



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 139402**

**TO: Janet Epps-Ford**  
**Location: REM-2C05/2C18**  
**Art Unit: 1635**  
**Thursday, December 02, 2004**  
**Case Serial Number: 09/551494**

**From: Paul Schulwitz**  
**Location: Biotech-Chem Library**  
**REM-1A65**  
**Phone: (571)272-2527**

**paul.schulwitz@uspto.gov**

### **Search Notes**

Examiner Epps-Ford,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Paul Schulwitz  
Technical Information Specialist  
STIC Biotech/Chem Library  
(571)272-2527

OSP 12/1

**Schulwitz, Paul**

---

**From:** Epps-Ford, Janet  
**Sent:** Wednesday, December 01, 2004 12:33 PM  
**To:** Schulwitz, Paul  
**Subject:** Sequence and Word-Search

Hi there, I was wondering if I can get the following claim language searched:

Application 09/551,494

An RNA viral vector comprising the nucleotide sequence of SEQ ID NO: 5 from the nucleotide at position 5430 to the nucleotide at position 5505.

*Thanks,*

*Janet L. Epps-Ford, Ph.D.*

*Art Unit 1635*

*Mailbox: Remsen 2C18*

*Office: Remsen 2C05*

*Phone: 571-272-0757*

*Fax: 571-273-0757*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: December 1, 2004, 20:12:59 ; Search time 408 Seconds  
(without alignments)  
977.833 Million cell updates/sec

Title: US-09-551-494-5\_COPY\_5430\_5505

Sequence: 1 gtgcagacggtcgcgaat.....tgaagtaaccaatgctgtga 76

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 4134886 segs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Geneseg 23Sep04:\*  
2: genesegm1980s:\*  
3: genesegm1990s:\*  
4: genesegm2000s:\*  
5: genesegm2001as:\*  
6: genesegm2001bs:\*  
7: genesegm2002as:\*  
8: genesegm2002bs:\*  
9: genesegm2003as:\*  
10: genesegm2003bs:\*  
11: genesegm2003ds:\*  
12: genesegm2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	100.0	6355	3 AAC62372	AAC62372 CDNA sequ
2	76	100.0	6355	9 ACC85005	ACC85005 TMV-U2 ge
3	69.6	91.6	769	10 ADI26338	ADI26338 Novel end
4	69.6	91.6	769	11 ADM68451	ADM68451 Tobacco m
5	69.6	91.6	769	12 ADP26604	ADP26604 Tobamovir
6	68	89.5	772	10 ADI26344	ADI26344 Novel end
7	68	89.5	772	11 ADM68457	ADM68457 Mosaic vi
8	58	89.5	772	12 ADP26610	ADP26610 Heterodup
9	58.6	77.1	411	3 AAC62379	AAC62379 Origin of
10	48.8	64.2	769	10 ADI26343	ADI26343 Novel end
11	48.8	64.2	769	11 ADM68456	ADM68456 Mosaic vi
12	48.8	64.2	769	12 ADP26609	ADP26609 Heterodup
13	45.8	60.3	5997	2 AAQ12188	AAQ12188 Odontoglo
14	45.8	60.3	6597	2 AAQ38106	AAQ38106 ORSV CDNA
15	44	57.9	356	1 AAQ30116	AAQ30116 TMV-RNA f
16	44	57.9	1019	1 AAQ30115	AAQ30115 TMV-RNA f
17	42.4	55.8	557	10 ADK56952	ADK56952 Plant DNA
18	42.4	55.8	557	10 ADK56952	ADK56952 Plant DNA
19	42.4	55.8	632	2 AA16647	AA16647 Tobacco m
20	42.4	55.8	732	10 ADI26342	ADI26342 Novel end
21	42.4	55.8	732	11 ADM68455	ADM68455 Mosaic vi

C	22	42.4	55.8	792	12	ADP26608	ADP26608 Heterodup
	23	42.4	55.8	804	12	ADN63297	ADN63297 Tobamovir
	24	42.4	55.8	805	6	ABT06582	ABT06582 Wild-type
	25	42.4	55.8	807	6	ABT06572	ABT06572 Wild-type
	26	42.4	55.8	807	6	ABX14577	ABX14577 Tobacco m
	27	42.4	55.8	807	11	ADM68440	ADM68440 Tobacco m
	28	42.4	55.8	807	12	ADP26593	ADP26593 Tobamovir
	29	42.4	55.8	895	2	AAQ62653	AAQ62653 Tobacco-m
	30	42.4	55.8	1187	10	ADD17932	ADD17932 DNA (Seqi
	31	42.4	55.8	1187	10	ADK56953	ADK56953 Plant DNA
	32	42.4	55.8	1825	2	AAQ67663	AAQ67663 TMV repl1
	33	42.4	55.8	1825	2	AAV54825	AAV54825 replicon
	34	42.4	55.8	1971	3	AAQ62382	AAQ62382 CDNA sequ
	35	42.4	55.8	5484	8	ADN15011	ADN15011 Tobacco m
	36	42.4	55.8	6395	2	AAQ95155	AAQ95155 Tobacco m
	37	42.4	55.8	6395	2	AAZ20642	AAZ20642 TMV-based
	38	42.4	55.8	6395	2	AAZ20642	AAZ20642 TMV-based
	39	42.4	55.8	6395	4	AAE82330	AAE82330 CDNA sequ
	40	42.4	55.8	6425	2	AAZ20644	AAZ20644 Wild-type
	41	42.4	55.8	6425	2	AAE82332	AAE82332 TMV-based
	42	42.4	55.8	6439	2	AAZ20643	AAZ20643 TMV-based
	43	42.4	55.8	6439	2	AAE82331	AAE82331 Tobacco m
	44	42.4	55.8	6446	2	AAZ20645	AAZ20645 TMV-based
	45	42.4	55.8	6446	4	AAE82334	AAE82334 Tobacco m

## ALIGNMENTS

RESULT 1	AAC62372	standard; DNA; 6355 BP.
ID	AAC62372	
AC	AAC62372;	
DT	19-MAR-2001	(first entry)
DE	CDNA sequence of the genome of tobacco mosaic virus-U2.	
XX	Plant phenotype; gene trait; Nicotiana; Oryza sativa; Zea mays; Brassica;	
KM	Gossypium; Triticum; Arabidopsis; Petunia; herbicide; transgenic plant;	
KM	tobacco necrosis virus; TMV; tobacco mosaic virus; TMV; helper virus; ss.	
OS	Tobacco mosaic virus.	
EN	MO200063397-A2.	
PD	26-OCT-2000.	
PF	17-APR-2000; 2000WO-EP003521.	
PR	20-APR-1999; 99US-00294022.	
XX	{AVET } AVENTIS CROSCIENCE NV.	
XX	Meulwaeeter F, Cornelissen M, Jacobs J, Van Eldik G, Metzlaiff M,	
XX	WPI; 2000-687182/67.	
PT	Identifying and isolating genes involved in determining the trait or	
PT	phenotype of plant species, by infecting plants with gene silencing	
PT	constructs targeted to the gene, and identifying plants with altered	
PT	traits.	
XX	Example 1; Page 53-56; 64pp; English.	
PS	The specification describes a method for isolating genes that determine a	
XX	trait or phenotype of a plant species. The method comprises identifying a	
CC	set of nucleic acids of genes correlated with the trait, creating a	
CC	library of gene silencing constructs in a viral RNA vector, targeting the	
CC	gene silencing constructs to the nucleic acid set, infecting a collection	
CC	of individual plants with these, identifying plants with altered traits	
CC	or phenotype, and isolating genes of the invention. The method is useful.	

CC for isolating genes involved in the determination of trait or a phenotype  
 CC of a plant such as Nicotiana, Oryza sativa, Zea mays, Brassica, Cossypium,  
 CC Triticum, Arabidopsis or Petunia. The method is also useful for  
 CC modulating the expression of selected nucleic acid sequences and for  
 CC validating the function of a nucleic acid sequence whose expression is  
 CC correlated with the presence or absence of a specific trait in plants,  
 CC but with otherwise unknown function. The method is also useful for  
 CC developing agronomically useful products such herbicides or transgenic  
 CC plants. The present sequence represents the cDNA sequence of the genome  
 CC of tobacco mosaic virus (TMV)-U2. The sequence was used to construct a  
 CC plasmid vector for the synthesis of an infective hybrid tobacco mosaic  
 CC virus (TMV)/satellite tobacco necrosis virus (STNV) helper virus RNA.  
 CC This helper virus is used in the method of the invention  
 CC  
 XX

Sequence 6355 BP; 1933 A; 1112 C; 1489 G; 1821 T; 0 U; 0 Other;

Query Match 100.0%; Score 76; DB 3; Length 6355;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-16;  
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGACAGACGGCTCGCCATTGAATCAGTGAAGGTTGTTAGAGATTGATGAA 60  
 Db 5430 GTGACAGACGGCTCGCCATTGAATCAGTGAAGGTTGTTAGAGATTGATGAA 5489

Qy 61 GTACCAATGGCTGTGA 76  
 Db 5490 GTACCAATGGCTGTGA 5505

## RESULT 2

ACC85005  
 ID ACC85005 standard; DNA; 6355 BP.

AC ACC85005;

DT 13-OCT-2003 (first entry)

XX TMV-U2 genome nucleotide sequence.

XX Inhibitory RNA; viral RNA vector; coat protein; TMV; U2; gene; ds.

XX Tobacco mosaic virus.

XX WO2003052108-A2.

XX 26-JUN-2003.

XX 05-DEC-2002; 2002WO-EP013964.

XX 18-DEC-2001; 2001US-0340488F.

XX (FARB) BAYER BIOSCIENCE NV.

XX Metzlaiff MH, Gosselle VML, Meulwaeter F, Fache ICA;

XX MPI; 2003-523529/49.

PT Introducing inhibitory RNA into a plant cell comprises providing a viral  
 PT RNA vector derived from a satellite RNA virus that encodes a coat  
 PT protein, and infecting a plant with the viral RNA vector and a  
 PT corresponding helper virus.

XX Example; Page 79-82; 86pp; English.

XX The invention relates to introducing inhibitory RNA into a plant cell.

XX The method involves providing a viral RNA vector derived from a satellite  
 CC RNA virus having a sequence that encodes a coat protein, and infecting a  
 CC plant with the viral RNA vector and a corresponding helper virus. The  
 CC methods and viral RNA vectors are useful in introducing inhibitory RNA  
 CC into plant cells. These may be used to determine or validate the function  
 CC of isolated nucleic acid sequences in plants. The present sequence  
 CC represents the nucleotide sequence of the genome of tobacco mosaic virus  
 CC (TMV)-U2

XX Sequence 6355 BP; 1933 A; 1112 C; 1489 G; 1821 T; 0 U; 0 Other;

Query Match 100.0%; Score 76; DB 9; Length 6355;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-16;  
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGACAGACGGCTCGCCATTGAATCAGTGAAGGTTGTTAGAGATTGATGAA 60  
 Db 5430 GTGACAGACGGCTCGCCATTGAATCAGTGAAGGTTGTTAGAGATTGATGAA 5489

Qy 61 GTACCAATGGCTGTGA 76  
 Db 5490 GTACCAATGGCTGTGA 5505

## RESULT 3

AD126338  
 ID AD126338 standard; DNA; 769 BP.

AC AD126338;

DT 22-APR-2004 (first entry)

XX Novel endonuclease Res I-related clone DNA 3.

XX endonuclease; molecular biology; plant propagation; phenotypic trait;

XX herbicide tolerance; heat tolerance; cold tolerance; drought; salinity;

XX osmotic stress; pest resistance; insect; nematode; arachnid; fungal;

XX bacterial; viral; enzyme production; secondary metabolite;

XX male sterility; female sterility; dwarfness; early maturity; Res I; ds.

XX Tobacco mosaic virus.

XX US2003148315-A1.

XX 07-AUG-2003.

XX 01-AUG-2002; 2002US-00211079.

XX 01-FEB-2002; 2002US-0353722P.

XX 14-MAR-2002; 2002US-00098155.

XX (PADG/) PADGETT H S.

XX (VAEM/) VAEMHONGS A A.

XX Padgett HS, Vaemhongs AA;

XX MPI; 2003-897548/82.

XX Example 15; Fig 8; 46pp; English.

CC This invention relates to a novel endonuclease (Res I) nucleic acid  
 CC molecule which comprises a fully defined sequence of 899 bp given in the  
 CC specification. The methods and compositions of the present invention are  
 CC useful in molecular biology, and more specifically to generating  
 CC populations of related nucleic acid molecules. They may also be used in  
 CC plant propagation with useful phenotypic traits, such as improved  
 CC tolerance to herbicides, improved tolerance to extremes of heat or cold,  
 CC drought, salinity or osmotic stress, improved resistance to pests  
 CC (insects, nematodes or arachnids) or diseases (fungal, bacterial or  
 CC viral), production of enzymes or secondary metabolites, male or female  
 CC sterility, dwarfness and early maturity. The present sequence is that of  
 CC a clone which was derived during the exemplification of the invention.

Sequence 769 BP; 247 A; 102 C; 201 G; 219 T; 0 U; 0 Other;

Query Match 91.6%; Score 69.6; DB 10; Length 769;  
 Best Local Similarity 94.7%; Pred. No. 2.3e-14;

Matches 72; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GTGACGACGGCTCCGCAATTGAACTCACTGAAAAAGTTGTAGAGATTCTGGATGAA 60  
DB 541 GTTACGACGAGCGCTCCGCAATTGAACTCACTGAAAAAGTTGTAGAGATTCTAGATGAA 600  
OY 61 GTACCAATGGCTGTGA 76  
DB 601 GTACCAATGGCTGTGA 616

## RESULT 4

ID ADM68451 standard; DNA; 769 BP.

AC ADM68451;

DT 03-JUN-2004 (first entry)

DE Tobacco mosaic virus movement protein gene #4.

KW ds; mismatch endonuclease; endonuclease; gene shuffling technology;

KM single nucleotide polymorphism; cancer susceptibility;

KV sequence variation redistribution; movement protein; gene.

OS Tobacco mosaic virus.

PN US2003157682-A1.

PD 21-AUG-2003.

PF 31-JAN-2003; 2003US-00356708.

PR 01-FEB-2002; 2002US-0353722P.

PR 14-MAR-2002; 2002US-00098155.

PR 01-AUG-2002; 2002US-00211079.

PA (PADG/) PADGETT H S.

PA (VAEH/) VAERHONGS A A.

PA (VOUD/) VOUDANT F S.

PA (SMIT/) SMITH M L.

PA (LIND/) LINDBO J A.

PA (FITZ/) FITZMAURICE W P.

XX Padgett HS, Vaerhongs AA, Vojdani FS, Smith ML, Lindbo JA;

XX Fitzmaurice WP;

XX WPI; 2003-766176/72.

XX Making a mismatch endonuclease, useful in gene shuffling and in detection

PT of single nucleotide polymorphisms, comprises transfecting a host with a

PT recombinant viral vector including a polynucleotide encoding a mismatch

PT endonuclease.

XX Example 14; SEQ ID NO 20; 79pp; English.

XX The invention relates to a method of making a mismatch endonuclease

XX enzyme comprising transfecting a host plant, animal, yeast, fungus or

XX bacterium with a recombinant viral vector that encodes a polynucleotide

XX sequence for a mismatch endonuclease, growing the host so that the

XX polynucleotide is expressed, and extracting the mismatch endonuclease

XX enzyme from the host. The method is useful for making mismatch

XX endonuclease enzymes, for obtaining peptides and polynucleotides with

XX desired functional properties and for detecting mutations. The mismatch

XX endonuclease enzymes are useful in gene shuffling technology for

XX developing new genes, in detecting single nucleotide polymorphisms for

XX e.g. detecting evidence of cancer susceptibility, or in redistributing

XX sequence variations between non-identical polynucleotide sequences. The

XX present sequence represents a tobacco mosaic virus movement protein gene.

XX Sequence 769 BP; 247 A; 102 C; 201 G; 219 T; 0 U; 0 Other;

XX Query Match 9.6%; Score 69.6; DB 11; Length 769;

Best Local Similarity 94.7%; Pred. No. 2,3e-14;

Matches 72; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GTGACGACGGCTCCGCAATTGAACTCACTGAAAAAGTTGTAGAGATTCTGGATGAA 60  
DB 541 GTTACGACGAGCGCTCCGCAATTGAACTCACTGAAAAAGTTGTAGAGATTCTAGATGAA 600  
OY 61 GTACCAATGGCTGTGA 76  
DB 601 GTACCAATGGCTGTGA 616

## RESULT 5

ID ADP26604 standard; DNA; 769 BP.

AC ADP26604;

DT 26-AUG-2004 (first entry)

DE Tobamovirus DNA #5.

KW Sequence variation; heteroduplex; transcription; DNA integration;

KM ribozyme expression; gene; ds; tobamovirus.

OS Tobamovirus.

PN US2004110130-A1.

PD 10-JUN-2004.

PF 25-OCT-2002; 2002US-00280913.

PR 02-FEB-2001; 2001US-026638EP.

PR 14-FEB-2001; 2001US-0268785P.

PR 01-FEB-2002; 2002US-0006635C.

PR 08-AUG-2002; 2002US-0402342P.

PA (LARG-) LARGE SCALE BIOLOGY CORP.

PA Padgett HS, Lindbo JA, Fitzmaurice WP;

XX WPI; 2004-440326/41.

XX Redistributing sequence variations between non-identical polynucleotide

XX sequences, useful for generating improved polynucleotide having a desired

XX characteristic, comprises making a heteroduplex and introducing a nick.

XX Example 15; SEQ ID NO 20; 75pp; English.

XX The invention relates to an in vitro method of redistributing sequence

XX variations between non-identical polynucleotide sequences, comprising

XX making a heteroduplex polynucleotide from two non-identical

XX polynucleotides, introducing a nick in the second strand at or near a

XX mismatch site where the nick occurred and using the first strand as a

XX template to replace the removed base(s) with bases that complement the

XX base(s) in the first strand. The invention also relates to an in vitro

XX method of making a population of sequence variants from a heteroduplex

XX polynucleotide sequence, obtaining a polynucleotide sequence encoding a

XX desired functional property and identifying a reassorted DNA molecule

XX encoding a protein with a desired functional property. The method is

XX useful for generating an improved polynucleotide sequence or a population

XX of improved polynucleotide sequences possessing at least one desired

XX phenotypic characteristic (e.g., promotes transcription of linked

XX polynucleotides), where such polynucleotides are useful for expression

XX from a plant, animal, fungal, yeast, or bacterial expression vector, for

XX integration to form a transgenic plant, animal or microorganism, and for

XX expression of a ribozyme. This sequence represents DNA used in the scope

XX of the invention.

XX Sequence 769 BP; 247 A; 102 C; 201 G; 219 T; 0 U; 0 Other;

XX Query Match 9.6%; Score 69.6; DB 11; Length 769;

Query Match 91.6%; Score 69.6; DB 12; Length 769;  
 Best Local Similarity 94.7%; Pred. No. 2.3e-14;  
 Matches 72; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTGACAGAGCGCTCGCAATTGAATCACTGAAAAGTTGTTGAGAGTTGTTGATGAA 60  
 |||  
 DB 541 GTAAACAGAGCGCTCGCAATTGAATCACTGAAAAGTTGTTGAGAGTTGTTGATGAA 600  
 |||

QY 61 GTACCAATGGCTGTGA 76  
 |||  
 DB 601 GTACCAATGGCTGTGA 616  
 |||

RESULT 6  
 AD126344/c  
 ID AD126344 standard; DNA; 772 BP.  
 XX  
 AC AD126344;

DT 22-APR-2004 (first entry)  
 DE Novel endonuclease Res I-related clone DNA 9.

XX endonuclease; molecular biology; plant propagation; phenotypic trait;  
 KW herbicide tolerance; heat tolerance; cold tolerance; drought; salinity;  
 KW osmotic stress; pest resistance; insect; nematode; arachnid; fungal;  
 KW bacterial; viral; enzyme production; secondary metabolite;  
 KW male sterility; female sterility; dwarfness; early maturity; Res I; ds.

OS Tobacco mosaic virus.  
 OS Tomato mosaic virus.

XX US2003148315-A1.

XX 07-AUG-2003.

XX 01-AUG-2002; 2002US-00211079.

XX 01-FEB-2002; 2002US-0353722P.  
 PR 14-MAR-2002; 2002US-00098155.

XX (PADG/) PADGETT H S.  
 PA (VAEM/) VAEMHONGS A A.

XX Padgett HS, Vaewhongs AA;

XX WPI; 2003-897548/82.

DR New nucleic acid molecule encoding endonucleases, useful in molecular  
 XX biology, specifically to generating populations of related nucleic acid  
 PT molecules, and in plant propagation with useful phenotypic traits.

XX Example 15; Fig 14; 46pp; English.

XX This invention relates to a novel endonuclease (Res I) nucleic acid  
 CC molecule which comprises a fully defined sequence of 899 bp given in the  
 CC specification. The methods and compositions of the present invention are  
 CC useful in molecular biology, and more specifically to generating  
 CC populations of related nucleic acid molecules. They may also be used in  
 CC plant propagation with useful phenotypic traits, such as improved  
 CC tolerance to herbicides, improved tolerance to extremes of heat or cold,  
 CC drought, salinity or osmotic stress, improved resistance to pests  
 CC (insects, nematodes or arachnids) or diseases (fungal, bacterial or  
 CC viral), production of enzymes or secondary metabolites, male or female  
 CC sterility, dwarfness and early maturity. The present sequence is that of  
 CC a clone which was derived during the exemplification of the invention.

XX Sequence 772 BP; 227 A; 196 C; 109 G; 240 T; 0 U; 0 Other;

Query Match 89.5%; Score 68; DB 10; Length 772;  
 Best Local Similarity 93.4%; Pred. No. 8.2e-14;  
 Matches 71; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTGACAGAGCGCTCGCAATTGAATCACTGAAAAGTTGTTGAGAGTTGTTGATGAA 60  
 |||  
 DB 232 GTAAACAGAGCGCTCGCAATTGAATCACTGAAAAGTTGTTGAGAGTTGTTGATGAA 173  
 |||

QY 61 GTACCAATGGCTGTGA 76  
 |||  
 DB 172 GTACCAATGGCTGTGA 157  
 |||

RESULT 7  
 ADM68457/c  
 ID ADM68457 standard; DNA; 772 BP.  
 XX  
 AC ADM68457;

DT 03-JUN-2004 (first entry)

DE Mosaic virus movement protein gene GRAMMR clone #11.

XX ds; mismatch endonuclease; endonuclease; gene shuffling technology;  
 KW single nucleotide polymorphism; cancer susceptibility;  
 KW sequence variation redistribution; movement protein; gene.

OS Tobacco mosaic virus.  
 OS Tomato mosaic virus.

XX US2003157682-A1.

XX 21-AUG-2003.

XX 31-JAN-2003; 2003US-00356708.

XX 01-FEB-2002; 2002US-0353722P.

XX 14-MAR-2002; 2002US-00098155.

XX 01-AUG-2002; 2002US-00211079.

XX (PADG/) PADGETT H S.  
 PA (VAEM/) VAEMHONGS A A.

XX (VOJD/) VOJDANI F S.  
 PA (SMIT/) SMITH M L.

XX (LIND/) LINDHO J A.  
 PA (FITZ/) FITZMAURICE W P.

XX Padgett HS, Vaewhongs AA, Vojdani FS, Smith ML, Lindho JA;

XX Fitzmaurice WP;

XX WPI; 2003-766176/72.

XX Making a mismatch endonuclease, useful in gene shuffling and in detection

XX of single nucleotide polymorphisms, comprises transfecting a host with a

XX recombinant viral vector including a polynucleotide encoding a mismatch

XX endonuclease.

XX Example 14; SEQ ID NO 26; 79pp; English.

XX The invention relates to a method of making a mismatch endonuclease  
 CC enzyme comprising transfecting a host plant, animal, yeast, fungus or  
 CC bacterium with a recombinant viral vector that encodes a polynucleotide  
 CC sequence for a mismatch endonuclease, growing the host so that the  
 CC polynucleotide is expressed, and extracting the mismatch endonuclease  
 CC enzyme from the host. The method is useful for making mismatch  
 CC endonuclease enzymes, for obtaining peptides and polynucleotides with  
 CC desired functional properties and for detecting mutations. The mismatch  
 CC endonuclease enzymes are useful in gene shuffling technology for  
 CC developing new genes, in detecting single nucleotide polymorphisms for  
 CC e.g. detecting evidence of cancer susceptibility, or in redistributing  
 CC sequence variations between non-identical polynucleotide sequences. The  
 CC present sequence represents a mosaic virus movement protein gene GRAMMR  
 CC clone.

XX Sequence 772 BP; 227 A; 196 C; 109 G; 240 T; 0 U; 0 Other;

Query Match 89.5%; Score 68; DB 11; Length 772;

	Best Local Similarity	93.4%	Pred. No. 8,2e-14:	
Matches	71;	Conservative	0;	Mismatches 5; Indels 0; Gaps 0;
QY	1	GTGACAGACGGCTGCCCAATTGA	CTCACTGCAAAAGTTGTGAGAGTTCGTGATGA	60
Db	232	GTAAACAGACGGCTGCCCAATTGA	CTCACTGCAAAAGTTGTGAGAGTTCATATATGA	173
QY	61	GTACCAATGGCTGTGA		76
Db	172	GTACCAATGGCTGTGA		157

RESULT 8  
ADP26610/c  
ID ADP26610 standard; DNA; 772 BP.

DT 26-AUG-2004 (first entry)

Heteroduplex DNA #15.

KM Sequence variation; heteroduplex; transcription; DNA integration  
KM ribozyme expression; gene; ds.

2N US2004110130-A1.

PD 10-JUN-2004

PF 25-OCT-2002; 2002US-00280913.

PR 02-FEB-2001; 2001US-0266386P.

PR 01-FEB-2002; 2002US-00066390.

XX  
XX  
XXXX / XXXXXX  
XXXX

XX	PI	Budget	HC	T	Indb	TR	Estimate

AA  
DB  
WPT: 2004-440326/41

PT Redistributing some

characteristic, comprises making a heteroduplex and introducing a nick.

PS Example 15; SEQ ID NO 26; 75pp; English.

The invention relates to an *in vitro* method of redistributing sequence variations between non-identical polynucleotide sequences, comprising making a heteroduplex polynucleotide from two non-identical polynucleotides, introducing a nick in the second strand at or near a base pair mismatch site, removing the mismatched base(s) from the mismatch site where the nick occurred and using the first strand as a template to replace the removed base(s) with bases that complement the base(s) in the first strand. The invention also relates to an *in vitro* method of making a population of sequence variants from a heteroduplex polynucleotide sequence, obtaining a polynucleotide sequence encoding a desired functional property and identifying a reassorted DNA molecule encoding a protein with a desired functional property. The method is useful for generating an improved polynucleotide sequence or a population of improved polynucleotide sequences possessing at least one desired phenotypic characteristic (e.g., promotes transcription of linked polynucleotides), where such polynucleotides are useful for expression from a plant, animal, fungal, yeast, or bacterial microorganism, for integration to form a transgenic plant, animal or microorganism, and for expression of a ribozyme. This sequence represents DNA used in the scope of the invention.

Sequence 772 BP; 227 A; 196 C; 109 G; 240 T; 0 U; 0 Other;

```

Query Match 98.5%; Score 68; DB 12; Length 772;
Best Local Similarity 93.4%; Pred. No. 8.2e-14;
Matches 71; Conservative 0; Mismatches 5; Indels 0; Gaps 0

Qy 1 GTGACAGACGGCTCCGCAATTGACTCACTGAAAAGTGTGTGAGAGATTCGTGATGAA 60
Db 232 GTACACAGACGGCTCCGCAATTGACTCACTGAAAAAGTGTGTGATGAGATTCATAGTAA 173
Qy 61 GTACCAATGGCTGTGA 76
Db 172 GTACCAATGGCTGTGA 157

RESULT 9
AAC62379
ID AAC62379 standard; DNA; 411 BP.
AC AAC62379;
XX AAC62379;
XX 19-MAR-2001 (first entry)
DT
DE Origin of assembly (OAS) of a tobacco mosaic virus (TMV)-U2.
XX
XX Plant phenotype; gene trait; Nicotiana; Oryza sativa; Zea mays; Brassica;
XX Gossypium; Triticum; Arabidopsis; Petunia; herbicide; transgenic plant;
XX tobacco necrosis virus; TMV; tobacco mosaic virus; TMV; helper virus;
XX origin of assembly; ss.
XX
XX Tobacco mosaic virus.
XX
XX WO200063397-A2.
XX
XX 26-OCT-2000.
XX
XX 17-APR-2000; 2000WO-EP003521.
XX
XX 20-APR-1999; 99US-00294022.
XX
XX (AVENTIS CROSCIENCE NV.
XX
XX Meulwaeater F, Cornelissen M, Jacobs J, Van Eldik G, Metzlaaff M;
XX WPI; 2000-687132/67.
XX
XX
XX Example 1; Page 63; 64pp; English.
XX
XX The specification describes a method for isolating genes that determine a
XX trait or phenotype of a plant species. The method comprises identifying a
XX set of nucleic acids of genes correlated with the trait, creating a
XX library of gene silencing constructs in a viral RNA vector, targeting the
XX gene silencing constructs to the nucleic acid set, infecting a collection
XX of individual plants with these, identifying plants with altered traits
XX or phenotype, and isolating genes of the invention. The method is useful
XX for isolating genes involved in the determination of trait or a phenotype
XX of a plant such as Nicotiana, Oryza sativa, Zea mays, Brassica, Gossypium,
XX Triticum, Arabidopsis or Petunia. The method is also useful for
XX modulating the expression of selected nucleic acid sequences and for
XX validating the function of a nucleic acid sequence whose expression is
XX correlated with the presence or absence of a specific trait in plants,
XX but with otherwise unknown function. The method is also useful for
XX developing agronomically useful products such herbicides or transgenic
XX plants. The present sequence is an origin of assembly (OAS) of a tobacco
XX mosaic virus (TMV)-U2. The sequence is used to construct infective hybrid
XX tobacco mosaic virus (TMV)/tobacco necrosis virus (TMV) vectors, for use
XX in the method of the invention
XX
XX Sequence 411 BP; 140 A; 70 C; 93 G; 108 T; 0 U; 0 Other;

```

SQ Sequence 411 BP; 140 A; 70 C; 93 G; 108 T; 0 U; 0 Other;

Query Match 77.1%; Score 58.6; DB 3; Length 411;  
 Best Local Similarity 33.8%; Pred. No. 1.4e-10;  
 Matches 61; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 12 CTCGCCAATTGACATGCTGAGAGTTGAGATGTAAGTACCAATGCG 71  
 DB 3 CTCGCCAATTGACATGCTGAGAGTTGAGATGTAAGTACCAATGCG 62

QY 72 TGTGA 76  
 DB 63 TGTGA 67

RESULT 10  
 ADI26343/c  
 ID ADI26343 standard; DNA; 769 BP.  
 AC ADI26343;  
 XX  
 DT 22-APR-2004 (first entry)  
 DE Novel endonuclease Res I-related clone DNA 8.  
 XX  
 KM endonuclease; molecular biology; plant propagation; phenotypic trait;  
 KM herbicide tolerance; heat tolerance; cold tolerance; drought; salinity;  
 KM osmotic stress; pest resistance; insect; nematode; arachnid; fungal;  
 KM bacterial; viral; enzyme production; secondary metabolite;  
 KM male sterility; female sterility; dwarfness; early maturity; Res I; ds.  
 XX  
 OS Tobacco mosaic virus.  
 OS Tomato mosaic virus.  
 XX  
 PN US2003148315-A1.  
 XX  
 PD 07-AUG-2003.  
 XX  
 PF 01-AUG-2002; 2002US-00211079.  
 XX  
 PR 01-FEB-2002; 2002US-0353722P.  
 PR 14-MAR-2002; 2002US-00098155.  
 XX  
 PA (PADGETT H S.  
 PA (VAEWONGS A A.  
 PA (VABW/)  
 PI Padgett HS, Vaeuwongs AA;  
 PI  
 DR WPI; 2003-897548/82.  
 XX  
 PT New nucleic acid molecule encoding endonucleases, useful in molecular  
 PT biology, specifically to generating populations of related nucleic acid  
 PT molecule, and in plant propagation with useful phenotypic traits.  
 XX  
 PS Example 15; Fig 13; 46pp; English.  
 XX  
 CC This invention relates to a novel endonuclease (Res I) nucleic acid  
 CC molecule which comprises a fully defined sequence of 899 bp given in the  
 CC specification. The methods and compositions of the present invention are  
 CC useful in molecular biology, and more specifically to generating  
 CC populations of related nucleic acid molecules. They may also be used in  
 CC plant propagation with useful phenotypic traits, such as improved  
 CC tolerance to herbicides, improved tolerance to extremes of heat or cold,  
 CC drought, salinity or osmotic stress, improved resistance to pests  
 CC (insects, nematodes or arachnids) or diseases (fungal, bacterial or  
 CC viral), production of enzymes and secondary metabolites, male or female  
 CC sterility, dwarfness and early maturity. The present sequence is that of  
 CC a clone which was derived during the exemplification of the invention.  
 CC  
 XX Sequence 769 BP; 222 A; 205 C; 96 G; 246 T; 0 U; 0 Other;

Query Match 64.2%; Score 48.8; DB 10; Length 769;  
 Best Local Similarity 77.6%; Pred. No. 4.6e-07;  
 Matches 59; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GTGACAGACGCTGCGCAATTGAATCACTGAAAAAGTTGTTGAGAGTTGATGATGAA 60  
 DB 229 GTGACAGATGAGAGACCATGGAACCTTCAGAAAGATGTTATGATGATGATGAA 170

QY 61 GTACCAATGGCTGTGA 76  
 DB 169 GTACCAATGGCTGTGA 154

RESULT 11  
 ADM68456/c  
 ID ADM68456 standard; DNA; 769 BP.  
 AC ADM68456;  
 XX  
 DT 03-JUN-2004 (first entry)  
 DE Mosaic virus movement protein gene GRAMR clone #10.  
 XX  
 KM ds; mismatch endonuclease; endonuclease; gene shuffling technology;  
 KM single nucleotide polymorphism; cancer susceptibility;  
 KM sequence variation redistribution; movement protein; gene.  
 XX  
 OS Tobacco mosaic virus.  
 OS Tomato mosaic virus.  
 XX  
 PN US2003157682-A1.  
 XX  
 PD 21-AUG-2003.  
 XX  
 PF 31-JAN-2003; 2003US-00356708.  
 XX  
 PR 01-FEB-2002; 2002US-0353722P.  
 PR 14-MAR-2002; 2002US-00098155.  
 PR 01-AUG-2002; 2002US-00211079.  
 XX  
 PA (PADGETT H S.  
 PA (VAEWONGS A A.  
 PA (VOUD/)  
 PA (SMIT/)  
 PA (LINDO J A.  
 PA (FITZ/)  
 PI Padgett HS, Vaeuwongs AA, Vojdani FS, Smith ML, Lindo JA;  
 PI Fitzmaurice WP;  
 PI  
 DR WPI; 2003-766176/72.  
 XX  
 PT Making a mismatch endonuclease, useful in gene shuffling and in detection  
 PT of single nucleotide polymorphisms, comprises transfecting a host with a  
 PT recombinant viral vector including a polynucleotide encoding a mismatch  
 PT endonuclease.  
 XX  
 PS Example 14; SEQ ID NO 25; 79pp; English.  
 XX  
 CC The invention relates to a method of making a mismatch endonuclease  
 CC enzyme comprising transfecting a host plant, animal, yeast, fungus or  
 CC bacterium with a recombinant viral vector that encodes a polynucleotide  
 CC sequence for a mismatch endonuclease, growing the host so that the  
 CC polynucleotide is expressed, and extracting the mismatch endonuclease  
 CC enzyme from the host. The method is useful for making mismatch  
 CC endonuclease enzymes, for obtaining peptides and polynucleotides with  
 CC desired functional properties and for detecting mutations. The mismatch  
 CC endonuclease enzymes are useful in gene shuffling technology for  
 CC developing new genes, in detecting single nucleotide polymorphisms for  
 CC e.g. detecting evidence of cancer susceptibility, or in redistributing  
 CC sequence variations between non-identical polynucleotide sequences. The  
 CC present sequence represents a mosaic virus movement protein gene GRAMR  
 CC clone.  
 XX  
 SQ Sequence 769 BP; 222 A; 205 C; 96 G; 246 T; 0 U; 0 Other;

Query Match 64.2%; Score 48.8; DB 11; Length 769;





```

AC AA038106;
XX
XX 05-JUL-1993 (first entry)
XX
XX ORSV cDNA.
XX
XX Odontoglossum ring spot virus; screen; transformation; ds.
XX
XX Odontoglossum ring spot virus.
XX
XX JP05030975-A.
XX
XX 09-FEB-1993.
XX
XX 26-JUL-1991; 91JP-00276075.
XX
XX 26-JUL-1991; 91JP-00276075.
XX
XX (NIOC ) NIPPON OIL KK.
XX
XX WPI; 1993-087957/11.
XX
XX cDNA of Odontoglossum ring-spot virus gene - useful as vector for plants
XX and probes for screening the virus.
XX
XX Claim 2; Page 11-28; 30pp; Japanese.
XX
XX RNA was isolated from Odontoglossum ring spot virus and used to prepare
XX cDNA. The DNA or its restriction fragments can be used to screen for ORSV
XX or to detect genes related to ORSV. Vectors contg. the DNA sequence can
XX be used to transform B. coli; Bacillus subtilis; Agrobacterium or plant
XX cells for prodn. of the recombinant 130K, 180K and 30K proteins of ORSV,
XX coat proteins or their peptide fragments
XX
XX Sequence 6597 BP; 1970 A; 1170 C; 1425 G; 2032 T; 0 U; 0 Other;
SQ
Query Match 60.3%; Score 45.8; DB 2; Length 6597;
Best Local Similarity 76.7%; Pred. No. 9.2e-06;
Matches 56; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 1 GTGACAGACGCGCTCGCCATTGAACCTGAAAGGTTGTGAGAGATTGCGATGAA 60
Db 5405 GTTACGAGAAAGAGGCGCCACCGAAGTACTGAGAGAGTTGTGATGAGAGAA 5464
QY 61 GTACCAATGGCTG 73
Db 5465 GTTCTATGGCTG 5477
RESULT 15
ID AAN30116 standard; RNA; 356 BP.
XX
XX AAN30116;
AC
XX
XX 25-MAR-2003 (revised)
DT 02-NOV-1992 (first entry)
XX
XX TMV-RNA fragment I originating at the capped 5' end of the viral RNA and
XX extending into the coat protein gene.
XX
XX RNA plant virus vector; tobacco mosaic virus; ss.
XX
XX Tobacco mosaic virus.
XX
XX Key Location/Qualifiers
XX 5
XX misc_feature /tag= e
XX /label= bp No. 5400
XX misc_feature /tag= a
XX /label= nucleation region
XX 61..235
XX CDS 61..235

```

```

FT /tag= c
FT /label= coat protein gene
FT misc_signal 236..320
FT /tag= b
FT /label= control region
FT misc_feature 348..356
FT /tag= d
FT /note= "site of cleavage by ribonuclease H"
XX
XX EP6553-A.
XX
XX 22-DEC-1982.
XX
XX 27-MAY-1981; 81US-00267539.
XX
XX 27-MAY-1981; 81US-00267539.
XX
XX (CANA ) NAT RES COUNCIL CANADA.
XX
XX Pelcher LE, Halasa MC;
XX
XX WPI; 1983-00323K/01.
XX
XX RNA plant virus vector from tobacco mosaic virus etc. - for modifying
XX genes in plants to alter growth disease resistance etc.
XX
XX Example; Page 27; 56pp; English.
XX
XX The inventors claim an RNA plant virus vector from tobacco mosaic virus.
XX The vector comprises a nucleotide sequence originating from the 5' end of
XX the (+) strand of the viral RNA (Fragment I) (see AAN30116) and a
XX sequence originating from the 3' end of the (+) strand (Fragment II) (see
XX AAN30114, AAN30115). The RNA vector may have foreign genetic information
XX inserted or attached, ultimately in the form of RNA, to the vector. The
XX mode of Fragment I and II production fragments enables fragments of any
XX desired length from any location in TMV-RNA to be generated (see
XX AAN30117, AAN30118). Fragment I 76 is designated Frag.I CP3. The viral
XX replicase gene is likely included Fragment I. (Updated on 25-MAR-2003 to
XX correct PA field.)
SQ
Sequence 356 BP; 121 A; 48 C; 87 G; 0 T; 100 U; 0 Other;
Query Match 57.9%; Score 44; DB 1; Length 356;
Best Local Similarity 56.6%; Pred. No. 1.8e-05;
Matches 43; Conservative 13; Mismatches 20; Indels 0; Gaps 0;
QY 1 GTGACAGACGCGCTCGCCATTGAACCTGAAAGGTTGTGAGAGATTGCGATGAA 60
Db 52 GUGAGAGACCGAGGCGCCCAUGGAACTUACAGAGAGAGUUGUGAUGCAUGAGAGAU 111
QY 61 GTACCAATGGCTGTA 76
Db 112 GUCCCUAUGGCAAUCA 127

```

Search completed: December 2, 2004, 07:53:21  
 Job time : 410 secs



QY 61 GTACCAATGGCTG 73  
 DB 62 GTTCCATGGCTG 50

RESULT 2  
 BM067518/c 348 bp mRNA linear EST 11-SEP-2002  
 LOCUS BM067518  
 DEFINITION KS08006E10 KS08 Capsicum annuum cDNA, mRNA sequence.  
 ACCESSION BM067518  
 VERSION BM067518.1 GI:22787638  
 KEYWORDS Capsicum annuum  
 SOURCE Capsicum annuum  
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Capsicum.

REFERENCE 1 (bases 1 to 348)  
 Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S., Hur, C.-G. and Choi, D.  
 Generation of Expressed Sequence Tags from Hot Pepper (Capsicum annuum L.) and Sequence Analysis in Relation to Hypersensitive Response Against Pathogen  
 Unpublished (2001)  
 CONTACT: Doil Choi  
 Genome Research Center and National Center for Genome Information  
 Korea Research Institute of Bioscience and Biotechnology  
 P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea  
 Tel: 82-42-860-4340  
 Fax: 82-42-860-4309  
 Email: doil@mail.kribb.re.kr  
 High quality sequence stop: 348.

FEATURES  
 source  
 1..348  
 Location/Qualifiers  
 /organism="Capsicum annuum"  
 /mol\_type="mRNA"  
 /cultivar="Hang Keun"  
 /db\_xref="taxon:4072"  
 /tissue\_type="anther"  
 /dev\_stage="10 weeks after germination"  
 /clone\_lib="KS08"  
 /note="Vector: pBluescript SK(-)"

ORIGIN  
 Query Match 51.8%; Score 39.4; DB 4; Length 348;  
 Best Local Similarity 71.2%; Pred. No. 0.011;  
 Matches 52; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 GTGACGAGCGCTCGCCATTGAACCTGAAAGGTTGTGAGAGTTCGTGAGTGA 60  
 DB 294 GTTCCGAGAGGAGGAGCCGTTGAATTACAGAGCAGTTGTATGATTCATCGAATCA 235

QY 61 GTACCAATGGCTG 73  
 DB 234 GTTCCATGGCTG 222

RESULT 3  
 CB264749 583 bp mRNA linear EST 06-NOV-2003  
 LOCUS 41-E014660-035-002-B11-T7R MP12-ADIS-035 Arabidopsis thaliana cDNA  
 DEFINITION clone MP12p2000B112Q 5-PRIME, mRNA sequence.  
 ACCESSION CB264749  
 VERSION CB264749.1 GI:32889522  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 583)  
 Schmid, R.J., Soerensen, T.R., Stracke, R., Torjek, O., Altman, T., Mitchell-Olds, T. and Weisshaar, B.

TITLE Large-scale identification and analysis of genome-wide single-nucleotide polymorphisms for mapping in Arabidopsis thaliana  
 JOURNAL Genome Res. 13 (6), 1250-1257 (2003)  
 MEDLINE 22683290  
 PUBMED 12793357

COMMENT  
 Contact: Weisshaar B  
 ADIS DNA core facility at MP12  
 Max-Planck-Institute for Plant Breeding Research  
 Carl-von-Linne Weg 10, 50829 Koeln, Germany  
 Fax: 00492215062851  
 Email: weisshaar@mp12-koeln.mpg.de  
 Insert Length: 583 Std Error: 0.00  
 Plate: 2 row: B column: 11  
 Seq primer: T7R: CTAATCGACTCACTATAGGGA.

FEATURES  
 source  
 1..583  
 Location/Qualifiers  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /cultivar="Achkarren-2 (Ax-2)"  
 /db\_xref="GABI:594666"  
 /db\_xref="taxon:3732"  
 /clone="MP12p2000B112Q"  
 /tissue\_type="inflorescence"  
 /adb\_host="E. coli TOP10"  
 /clone\_lib="MP12-ADIS-035"  
 /note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; cDNA library from Arabidopsis thaliana, accession Achkarren-2; inflorescences from flower buds to young siliques; library was made at the Max-Planck-Institute for Plant Breeding Research, Cologne, Germany; cloning sites SalI-NotI; primer sites and orientation:  
 T7-SalI-CCACGCGTCCG-5prime-cDNA-polyA-CC-NotI-SP6; Note: Sequencing granted in the context of the GABI Arabidopsis Verbund: Genetic Diversity, Establishment of high-efficiency SNP-based mapping tools and development of methods for genome-wide mutation detection" PI: Bernd Weisshaar; Sequence submission managed by RZPD/GABI-Primary database: http://gabi.rzpd.de This clone is available from RZPD; contact RZPD (clone@rzpd.de) for further information."

ORIGIN  
 Query Match 42.9%; Score 32.6; DB 6; Length 583;  
 Best Local Similarity 66.2%; Pred. No. 2.6;  
 Matches 47; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 6 AGACGGCTCGCCATTGAACCTGAAAGGTTGTGAGAGTTCGTGAGTGAAGTACC 65  
 DB 381 AGATGCTTCCCACTTAATCATGATGAATGCTCTTGAGAAATTGGAGCAATATCC 440

QY 66 AATGGCTGTGA 76  
 DB 441 AATGCTTCA 451

RESULT 4  
 B12288 954 bp DNA linear GSS 14-MAY-1997  
 LOCUS T2M2-SP6 TAMU Arabidopsis thaliana genomic clone T2M2, genomic survey sequence.  
 ACCESSION B12288  
 VERSION B12288.1 GI:2093409  
 KEYWORDS GSS.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 954)  
 Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sur, H. and Ecker, J.  
 BAC End Sequences at ATGC  
 JOURNAL Unpublished (1997)

COMMENT  
Other GSSs: T2M2-T7  
Contact: Ecker J.  
Arabidopsis Thaliana Genome Center  
University of Pennsylvania  
Dept. of Biology, University of Pennsylvania, Philadelphia, PA  
19104  
Tel: 215-898-9384  
Fax: 215-898-8780  
Email: jecker@atgenomc.bio.upenn.edu  
Seq primer: Sp6  
Class: BAC ends  
High quality sequence start: 93  
High quality sequence stop: 163.  
Location/Qualifiers  
1. .954  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="T2M2"  
/sex="thermaproditic"  
/note="Vector: BelosACII; Site\_1: HindIII; Site\_2:  
HindIII; Produced by Rod Wing"

ORIGIN  
Query Match 41.6%; Score 31.6; DB 8; Length 954;  
Best Local Similarity 64.8%; Pred. No. 6.3;  
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
QY 6 AGACGGCTGGCAATGAACTGAAAGCTTGTGAGAGCTTCGTGATGAAGTACC 65  
DB 72 AATAGCTCTCCACTTGAATATGATGATTAAGCTTGTGAGAGTTCGAGAACCAATTCG 131  
QY 66 AATGCTGTGA 76  
DB 132 AATTGCTTTA 142

RESULT 5  
AV441961 389 bp mRNA linear EST 18-FEB-2004  
LOCUS AV441961 Arabidopsis thaliana above-ground organ two to six-week  
DEFINITION AV441961 Arabidopsis thaliana cDNA clone APD19F09\_1\_5, mRNA sequence.  
ACCESSION AV441961.1 GI:7612355  
VERSION EST.  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 389)  
Asamizu, E., Nakamura, Y., Sato, S., and Tabata, S.  
A large scale analysis of cDNA in Arabidopsis thaliana: Generation  
of 12,028 non-redundant expressed sequence tags from normalized and  
size-selected cDNA libraries  
DNA Res. 7 (3), 175-180 (2000)  
20363093  
10907847  
PUBMED  
COMMENT  
Contact: Erika Asamizu  
The First Laboratory for Plant Gene Research  
Kazusa DNA Research Institute  
Yana 1532-3, Kisarazu, Chiba 252-0812, Japan  
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/  
Location/Qualifiers  
1. .389  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/clone="APD19F09\_1\_5"  
/tissue\_type="aboveground organs"

FEATURES  
source

ORIGIN  
Query Match 40.8%; Score 31; DB 1; Length 389;  
Best Local Similarity 64.8%; Pred. No. 8.5;  
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
QY 6 AGACGGCTGGCAATGAACTGAAAGCTTGTGAGAGCTTCGTGATGAAGTACC 65  
DB 62 AATGCTCTCCACTTGAATATGATGATTAAGCTTGTGAGAGTTCGAGAACCAATTCG 121  
QY 66 AATGCTGTGA 76  
DB 122 AATTGCTTTA 132

RESULT 6  
CA963974 530 bp mRNA linear EST 03-JAN-2003  
LOCUS CAT1008D10AF Infected Arabidopsis Leaf Arabidopsis thaliana cDNA,  
DEFINITION mRNA sequence.  
ACCESSION CA963974.1 GI:27490531  
VERSION EST.  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 530)  
Lundgaard, M., Emmersen, J., Nielsen, K.L., Wilson, I., Somerville, S.  
and Welinder, K.G.  
EST sequencing of Erysiphe cichoracearum infected Arabidopsis  
plants  
Unpublished (2002)  
Contact: Karen G. Welinder  
Institut for bioteknologi  
Aalborg Universitet  
Schougardsholmsvej 49, 9000 Aalborg, Denmark  
Tel: +45 96358467  
Fax: +45 98141808  
Email: kywebio.auc.dk.  
Location/Qualifiers  
1. .530  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/ecotype="Columbia"  
/db\_xref="taxon:3702"  
/dev stage="Plant 3 weeks old, three days post infection"  
/note="Organ: Leaf; Vector: pBluescript; Mixed cDNA  
library of Arabidopsis and E. cichoracearum infected leaf  
from three weeks old Arabidopsis plants. Plants were  
harvested 3 days after infection and mRNA oligo dt  
selected."

ORIGIN  
Query Match 40.8%; Score 31; DB 6; Length 530;  
Best Local Similarity 64.8%; Pred. No. 9;  
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
QY 6 AGACGGCTGGCAATGAACTGAAAGCTTGTGAGAGCTTCGTGATGAAGTACC 65  
DB 215 AATGCTCTCCACTTGAATATGATGATTAAGCTTGTGAGAGTTCGAGAACCAATTCG 274  
QY 66 AATGCTGTGA 76  
DB 275 AATTGCTTTA 285

```

RESULT 7
LOCUS      FR020421
DEFINITION F.rubripes GSS sequence, clone 041p11b11, genomic survey sequence.
ACCESSION  AL013304
VERSION     AL013304.1 GI:2679672
KEYWORDS   GSS; genome survey sequence.
SOURCE     Takifugu rubripes (Fugu rubripes)
ORGANISM   Takifugu rubripes

REFERENCE
AUTHORS    Edgar, G., Clark, M.S., Meek, S., Smith, S., Warner, S., Edwards, Y.J.,
            Bouchere, N., Cottage, A., Yeo, G.S., Umanian, Y., Williams, G. and
            Brenner, S.
TITLE       Generation and analysis of 25 Mb of genomic DNA from the pufferfish
            Fugu rubripes by sequence scanning
JOURNAL     Genome Res. 9 (10), 960-971 (1999)
MEDLINE     99455097
PUBMED      10523524
REFERENCE   2 (bases 1 to 591)
AUTHORS     Edgar, G., Clark, M., Smith, S., Meek, S., Warner, S., Umanian, Y.,
            Williams, G. and Brenner, S.
TITLE       Direct Submision
JOURNAL     Submitted (08-DEC-1997) MRC Human Genome Mapping Project Resource
            Centre Hinxton, Cambridge, CB10 1SB, E-mail: biolhelp@hmp.mrc.ac.uk
COMMENT     V type: phagemid
            PRIMER: KS
            DESC: One pass dye-terminator sequencing of cosmid cloned genomic
            sequence.
FEATURES
source      location/Qualifiers
            1..591
               /organism="Takifugu rubripes"
               /mol_type="genomic DNA"
               /db_xref="taxon:31033"
               /clone="041p11b11"
               /clone_1lb="cosmid 041p11"

ORIGIN
Query Match      40.8%; Score 31; DB 9; Length 591;
Best Local Similarity 64.2%; Pred. No. 9.2;
Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Cy 10 GGCTCGCCAACTGACCTGAAAGGTTGGAGAGTTCGCGATGAAGTACCAATG 69
    |||
Db 113 GGCGTCGAACTGAACTGGAAGATGTTGACATCTTGGGAGCAACTGTAAAG 172
    |||

Qy 70 GCTGTGA 76
    |||
Db 173 GTGTGA 179

RESULT 8
LOCUS      BU636714
DEFINITION 010F08 infected Arabidopsis leaf Arabidopsis thaliana cDNA, mRNA
ACCESSION  BU636714
VERSION     BU636714.1 GI:23303969
KEYWORDS   EST.
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
            1 (bases 1 to 656)
            Lundsgaard, M., Emmersen, J., Nielsen, K.L., Wilson, I., Somerville, S.
            and Welinder, K.G.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source      location/Qualifiers
            1..656
               /organism="Arabidopsis thaliana"
               /mol_type="mRNA"
               /ecolyt="Columbia"
               /db_xref="taxon:3702"
               /dev_stage="Plant 3 weeks old, three days post infection"
               /clone_1lb="infected Arabidopsis leaf"
               /note="Organ: leaf; Vector: pBluescript; Mixed cDNA
               library of Arabidopsis and E. cichoracearum infected leaf
               harvested 3 days after infection and mRNA oligo dt
               selected."

ORIGIN
Query Match      40.8%; Score 31; DB 5; Length 656;
Best Local Similarity 64.8%; Pred. No. 9.4;
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Cy 6 AGACGGCTCGCCAACTGACCTGAAAGGTTGGAGAGTTCGCGATGAAGTACC 65
    |||
Db 363 AGATGCTTCCTCCACTGAAATCATGATTAAGCTTTGAGATCGGACCAATGCG 422
    |||

Qy 66 AATGCTGTGA 76
    |||
Db 423 AATGCTTTTA 433

```

```

TITLE       EST sequencing of Erysiphe cichoracearum infected Arabidopsis
            plants
JOURNAL     Unpublished (2002)
COMMENT     Contact: Karen G. Welinder
            Institut for bioteknologi
            Aalborg Universitet
            Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
            Tel: +45 96358467
            Fax: +45 98141808
            Email: kgw@bio.auc.dk.
FEATURES
source      location/Qualifiers
            1..656
               /organism="Arabidopsis thaliana"
               /mol_type="mRNA"
               /ecolyt="Columbia"
               /db_xref="taxon:3702"
               /dev_stage="Plant 3 weeks old, three days post infection"
               /clone_1lb="infected Arabidopsis leaf"
               /note="Organ: leaf; Vector: pBluescript; Mixed cDNA
               library of Arabidopsis and E. cichoracearum infected leaf
               harvested 3 days after infection and mRNA oligo dt
               selected."

ORIGIN
Query Match      40.8%; Score 31; DB 5; Length 656;
Best Local Similarity 64.8%; Pred. No. 9.4;
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Cy 6 AGACGGCTCGCCAACTGACCTGAAAGGTTGGAGAGTTCGCGATGAAGTACC 65
    |||
Db 363 AGATGCTTCCTCCACTGAAATCATGATTAAGCTTTGAGATCGGACCAATGCG 422
    |||

Qy 66 AATGCTGTGA 76
    |||
Db 423 AATGCTTTTA 433

```

## ORIGIN

Library of Arabidopsis and B. cinerosearum infected leaf  
from three weeks old Arabidopsis plants. Plants were  
harvested 3 days after infection and mRNA oligo dt  
selected."

Query Match 40.8%; Score 31; DB 5; Length 686;  
Best Local Similarity 64.8%; Pred. No. 9.5;  
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 6 AGAGCGCTGCCAATGAACTGAAAGGTTGAGAGTTGATGAAATACC 65  
DB 347 AGATGCTTCTCCACTGAAATCATGATGATTAAGCTCTTGAGATTGAGACCAATCGC 406

QY 66 AATGCTGTGA 76  
DB 407 AATGCTTTTA 417

## RESULT 10

CNSOADUC 1535 bp mRNA linear HTC 06-FEB-2004  
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
DEFINITION GSUTFB472C07 of Flowers and buds of strain col-0 of Arabidopsis

ACCESSION BX813913.1 GI:42474399  
VERSION HTG; GSUT\_cDNA.  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana (thale cress)

REFERENCE 1 (bases 1 to 1535)  
AUTHORS Castell, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,  
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,  
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.  
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1535)  
AUTHORS Genoscope.  
TITLE Direct Submission.  
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

COMMENT The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction; Temple G.  
Genoscope members carried out sequencing and annotation; Castell  
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,  
Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
length

http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
Location/Qualifiers  
1.1535  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSUTFB472C07"  
/tissue\_type="flowers and buds"  
/plasmid="pCMVSPORT 6"  
/complement(1..1535)  
/gene="At1g62180"

## FEATURES

source

gene

ORIGIN

Query Match 40.8%; Score 31; DB 3; Length 1535;  
Best Local Similarity 64.8%; Pred. No. 11;  
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 6 AGAGCGCTGCCAATGAACTGAAAGGTTGAGAGTTGATGAAATACC 65  
DB 305 AGATGCTTCTCCACTGAAATCATGATGATTAAGCTCTTGAGATTGAGACCAATCGC 364

QY 66 AATGCTGTGA 76  
DB 365 AATGCTTTTA 375

## RESULT 11

CNSOACVY 1579 bp mRNA linear HTC 06-FEB-2004  
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone  
DEFINITION GSUTL682B02 of Adult vegetative tissue of strain col-0 of  
Arabidopsis thaliana (thale cress).

ACCESSION BX815501.1 GI:42473038  
VERSION HTG; GSUT\_cDNA.  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana (thale cress)

REFERENCE 1 (bases 1 to 1579)  
AUTHORS Castell, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,  
Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,  
Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.  
TITLE Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:  
A Combined Approach to Evaluate and Improve Arabidopsis Genome  
Annotation

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1579)  
AUTHORS Genoscope.  
TITLE Direct Submission.  
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)

COMMENT The sequences are based on single pass reads.  
Life Technologies (a division of Invitrogen) members carried out  
full-length libraries construction; Temple G.  
Genoscope members carried out sequencing and annotation; Castell  
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,  
Schachter V., Weissenbach J., Salanoubat M.  
URGV INRA : Clepet C., Caboche M.  
Annotation is based on the June 2003 version of the Arabidopsis  
genome released by MIPS (Munich Information center for Protein  
Sequences). 5 prime and 3 prime are assembled with Phrap.  
http://www.genoscope.cns.fr/externe/sequences/Banque\_Projet\_EF/Full  
length

http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.  
Location/Qualifiers  
1.1579  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="Col-0"  
/db\_xref="taxon:3702"  
/clone="GSUTL682B02"  
/tissue\_type="Adult vegetative tissue"  
/plasmid="pCMVSPORT 6"  
/complement(1..1579)  
/gene="At1g62180"

## FEATURES

source

gene

ORIGIN

Query Match 40.8%; Score 31; DB 3; Length 1579;  
Best Local Similarity 64.8%; Pred. No. 11;  
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 6 AGAGCGCTGCCAATGAACTGAAAGGTTGAGAGTTGATGAAATACC 65

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: December 1, 2004, 20:23:09 ; Search time 1635 seconds  
(without alignments)  
2198.176 Million cell updates/sec

Title: US-09-551-494-5\_COPY\_5430\_5505

Perfect score: 76  
Sequence: 1 gtgacagacgctcgcacat.....tgaagaccacatgctcgtga 76

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 4526723 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: GenBank:  
1: gb\_ba:\*  
2: gb\_hg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_pl:\*  
8: gb\_pr:\*  
9: gb\_pt:\*  
10: gb\_rb:\*  
11: gb\_sg:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	100.0	6355	6 AX040177	AX040177 Sequence
2	76	100.0	6355	6 AX795380	AX795380 Sequence
3	76	100.0	6355	14 TMGCC	M34077 Tobacco mil
4	71.2	93.7	1413	14 TMGCC	M34236 Tobacco mil
5	69.6	91.6	6356	14 AB078435	AB078435 Tobacco m
6	58.6	77.1	411	6 AX040184	AX040184 Sequence
7	45.8	60.3	912	14 CRVTRANS	X55236 Odontogloss
8	45.8	60.3	5997	6 E03624	E03624 DNA encodin
9	45.8	60.3	6597	6 E04305	E04305 cDNA encodi
10	45.8	60.3	6609	14 ORU34586	U34586 Odontogloss
11	45.8	60.3	6611	14 S83257	S83257 126 kda pro
12	45.8	60.3	6612	14 AY571290	AY571290 Odontoglo
13	44	57.9	1004	14 TOTW2	V01406 TMV 3' end.
14	44	57.9	1019	6 E00088	E00088 Genomic RNA
15	44	57.9	6395	14 D63809	D63809 Tobacco mos
16	42.4	55.8	396	14 TM0307579	TM0307579 Tobacco m
17	42.4	55.8	632	6 A68760	A68760 Sequence 1
18	42.4	55.8	706	14 TM0509869	TM0509869 Tobacco m
19	42.4	55.8	750	14 TM0509080	TM0509080 Tobacco m

20	42.4	55.8	750	14 TM0509084	AJ509084 Tobacco m
21	42.4	55.8	782	14 TM0308692	AJ308692 Tobacco m
22	42.4	55.8	788	14 TM0308685	AJ308685 Tobacco m
23	42.4	55.8	790	14 TM0307583	AJ307583 Tobacco m
24	42.4	55.8	804	6 CQ793025	CQ793025 Sequence
25	42.4	55.8	804	14 TM0307582	AJ307582 Tobacco m
26	42.4	55.8	806	14 TM0308682	AJ308682 Tobacco m
27	42.4	55.8	806	14 TM0308684	AJ308684 Tobacco m
28	42.4	55.8	807	6 BD263879	BD263879 Viral exp
29	42.4	55.8	807	6 BD263880	BD263880 Viral exp
30	42.4	55.8	807	6 AR435699	AR435699 Sequence
31	42.4	55.8	807	6 AR435700	AR435700 Sequence
32	42.4	55.8	807	6 AX045758	AX045758 Sequence
33	42.4	55.8	807	6 AX592974	AX592974 Sequence
34	42.4	55.8	807	14 TM0307581	AJ307581 Tobacco m
35	42.4	55.8	807	14 TM0308683	AJ308683 Tobacco m
36	42.4	55.8	807	14 TM0308693	AJ308693 Tobacco m
37	42.4	55.8	807	14 TM0509081	AJ509081 Tobacco m
38	42.4	55.8	808	14 TM0307578	AJ307578 Tobacco m
39	42.4	55.8	808	14 TM0308686	AJ308686 Tobacco m
40	42.4	55.8	808	14 TM0308688	AJ308688 Tobacco m
41	42.4	55.8	808	14 TM0308690	AJ308690 Tobacco m
42	42.4	55.8	809	14 TM0308691	AJ308691 Tobacco m
43	42.4	55.8	811	14 TM0509082	AJ509082 Tobacco m
44	42.4	55.8	815	14 AY300161	AY300161 Tobacco m
45	42.4	55.8	817	14 TM0509083	AJ509083 Tobacco m

## ALIGNMENTS

RESULT 1  
AX040177  
LOCUS AX040177 6355 bp DNA linear PAT 18-NOV-2000  
DEFINITION Sequence 5 from Patent WO0063397.  
ACCESSION AX040177  
VERSION AX040177.1 GI:11230127  
KEYWORDS  
SOURCE  
SCORCE  
ORGANISM  
synthetic construct  
synthetic construct  
artificial sequences.

REFERENCE  
1 Meunier, F., Corneliisse, Jacobs, J., van Eijk, G. and Metzlaaff, M.  
METHODS AND MEANS FOR DELIVERING INHIBITORY RNA TO PLANTS AND APPL  
TITLES  
JOURNAL  
Patent: WO 0063397-A 5 26-OCT-2000;  
Aventis CropScience N.V. (BE)  
FEATURES  
source  
location/Qualifiers  
1..6355  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="cDNA copy of the nucleotide sequence of the genome  
of TMV-U2"

## ORIGIN

Query Match 100.0%; Score 76; DB 6; Length 6355;  
Best Local Similarity 100.0%; Pred. No. 2.2e-15;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACAGACGCTCGCCCAATGGAATCACTGAAAAAGTTGTGAGAGATTCGTGATGGA 60  
DB 5430 GTGACAGACGCTCGCCCAATGGAATCACTGAAAAAGTTGTGAGAGATTCGTGATGGA 5489  
QY 61 GTACCAATGAGCTGTGA 76  
DB 5483 GTACCAATGAGCTGTGA 5505

RESULT 2  
LOCUS AX795380 6355 bp mRNA linear PAT 04-OCT-2003  
DEFINITION Sequence 7 from Patent WO03052108.



```

ACCESSION   AX795380
VERSION     AX795380.1 GI:37516053
KEYWORDS
SOURCE      synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Metzlaff, M.H., Gossels, V.M., Neulwetter, F. and Fache, I.C.
  TITLE     Improved methods and means for delivering inhibitory rna to plants
            and applications thereof
  JOURNAL   Patent: WO 03052108-A 7 26 JUN-2003;
            Bayer Bioscience N.V. (BE)
FEATURES
  Source
    1..6355
    /organism="synthetic construct"
    /mol_type="mRNA"
    /db_xref="taxon:32630"
    /note="cDNA sequence of the genome of TMV-U2"

ORIGIN
Query Match      100.0%; Score 76; DB 6; Length 6355;
Best Local Similarity 100.0%; Pred. No. 2,2e-15;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2y      1 GTGCAAGCGCTGCGCAATTGAACCTGAGAAAGTTGTGAGAGACTTGTGATGAA 63
Db      5430 GTGCAAGCGCTGCGCAATTGAACCTGAGAAAGTTGTGAGAGACTTGTGATGAA 5489

Qy      61 GTACCAATGGCTGTGA 76
Db      5490 GTACCAATGGCTGTGA 5505

RESULT 3
TMGC      6355 bp ss-RNA linear VRL 03-AUG-1993
LOCUS      Tobacco mild green mosaic virus complete genome.
ACCESSION   M34077.1 GI:335243
VERSION     M34077.1 GI:335243
KEYWORDS
SOURCE      coat protein.
            Tobacco mild green mosaic virus
            Tobacco mild green mosaic virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
ORGANISM
REFERENCE   1 (bases 6127 to 6355)
  AUTHORS   Garcia-Arenal, F.
  TITLE     Sequence and structure at the genome 3' end of the U2-strain of
            tobacco mosaic virus, a histidine-accepting tobamovirus
  JOURNAL   Virology 167 (1), 201-206 (1998)
  MEDLINE   89045644
  PUBMED    3188396
REFERENCE   2 (bases 1 to 6355)
  AUTHORS   Solis, I. and Garcia-Arenal, F.
  TITLE     The complete nucleotide sequence of the genomic RNA of the
            tobamovirus tobacco mild green mosaic virus
  JOURNAL   Virology 177 (2), 553-558 (1993)
  MEDLINE   90320127
  PUBMED    2371769
COMMENT
  Original source text: Tobacco mild green mosaic virus (strain
  U2-TMV), cDNA to viral RNA, from N. tabacum cv. Samsum.
  Draft entry and computer-readable sequence for [1] kindly submitted
  by F. Garcia-Arenal, 10-FEB-1989. The RNA appears to have a tRNA-
  like, L-shaped structure at the 3' terminus, linked to a quasi-
  continuous double-helical stalk, with five pseudoknots involved in
  the formation of the whole structure. However, the structure of
  U2-TMV RNA is less stringently conserved than the 3' terminus of
  'vulgar' and other histidine-accepting tobamoviruses. Draft entry
  and computer-readable sequence for [1] kindly submitted by
  F. Garcia-Arenal, 08-MAY-1990, for release after publication.
  Location/Qualifiers
    1..6355
    /organism="Tobacco mild green mosaic virus"
    /mol_type="genomic RNA"
    /db_xref="taxon:12241"

```

```

CDS
join(71..3403,3407..4900)
/note="183 kDa protein"
/codon_start=1
/db_xref="GI:335243"
/translation="MAHIGSIISNALLSEVSGKNTLVNDLARRMYDTAAVEFNARD
RPNKFSKTISEBOTLVSNAYPEFOITFTYNOAIVSLAGLALRELYMLQVPG
SPYDIGNFPAH.FKGRDYHACMPNLDIPDIRHEQDQSIEMYLRSKNVLP
EPQBARIRYAAENVECCSKTPDCCIRHPENGRARYALASHLVDLPVHGAL
SKNHYVCAASILAELLDDQTEVTLNIGVTPKRGDVSFFFADESTLNSHYKN
ILAHYVCSYPPASSRIYVFKFPLVTVWVTCCKPTKDYTLIYVSRVGDSDPYE
AMEDAFAYKTLAMFENTERAIFRDTASVNFPPKDMVILPLGSGSTSKMRESEV
IYNDFYVTVLNIHRTYQAKALTYQNVLSFESLRSRYINGTVARSMDCKALLOP
LSMTPLOITKLALLODDIVMGKFCCLKTSSSELIMDEGKFGNVEPTXRLYSRXI
LDSENMALKIPIPLVYTMORPFAETTSSELPHLDIKDLBEACQYDALSELSTL
KGADNPDIAPKDMCKALDYSPIVAAYIVAENRSGTLTFDPFTEBNVAKAKST
ASEAVNCLBPTSEBVNANKSIAKGRLPVCAESHGLTNALREQLSEINDFKACV
DSVYTKQMASVYVYSLKVOQKNVYVSLAASLATAVSNLCKSLDEGVSDGSRKY
GWDVTLKWLKLPAAKGSWGVLDKGMKFTALLSEGRMATESPWRVAVSDT
MYSDIAKGLNLRKMRDGPHEPTAIVGVGGKGYMTEPFLDSDLLVPG
KQAAAMRRRANSGLIRATMDVRYTDSLIMKPKSHRLLPIDEGLMHTGCNPL
VLISGCIATYIGDQOIFPTIRYONPYPHREKLDQVEEMKRTILRCGDNVFL
QSKREGAVTTSTVQSRVSSEMIGKVLNVSXPKLQVTFQADPFLKEEGYANL
VNTVE-OGSETFEDVSLVRLATPLT.ISSSPHVALVLTWTATYSFKYTVLPLVQ
IISLSISSEFLMEYVEAGSR"
<71..3403
/note="183 kDa protein"
3407..4900
/note="183 kDa protein"
4890..5660
/note="28.5 kDa protein"
/codon_start=1
/protein_id="AAA47936.1"
/db_xref="GI:335246"
/translation="MAVSLDQTYKISFPIDSKDDELLPAMTYKSVRISTVDKIMA
WKDLSLDVJLKGVLVKRGYGLADLVGGENLDPNCKGVSVCIVKMRKRS
ALDGAHYAPACKKNFSFKLIPNYSITSEDAKHMQVLYN.KGVAEBGCPFLSEY
SICVYHKNRYKGLRERILLSTVDGSPLEIEKVEEFVEPMAVVKLEKVENKEMV
GNVNNKINNSGKGFKEIEIDNVSDDSIASSSF"
5666..6145

```

CDS

/note="coat protein"  
/codon\_start=1  
/protein\_id="AAA47937.1"  
/db\_xref="GI:335247"  
/translation="MPYINSPOFVYLSAYADPVOLINLCTNALGNOFOTQOART  
VQOQZALDMKVPSPMTVRFPSADFYVRYNLTDLPLTALLNSPTTRKLTLEVDQPA  
PMTTEYNATORVDATVAIRASININLANELVGTGEMQAGETASGLVWTTTPAT"

Query Match 100.0%; Score 76; DB 14; Length 6355;  
Best Local Similarity 100.0%; Pred. No. 2,2e-15;  
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACAGACGGCTGGCCAAATTGAACCTCACTGAAAAAGTTGTGAGAGCTTCGTGATGAA 60  
DB 5430 GTGACAGACGGCTGGCCAAATTGAACCTCACTGAAAAAGTTGTGAGAGCTTCGTGATGAA 5489  
QY 61 GTACCAATGGCTGTGA 76  
DB 5430 GTACCAATGGCTGTGA 5505

RESULT 4  
TMGMPCPA  
LOCUS  
DEFINITION  
Tobacco mild green mosaic virus movement and coat protein genes,  
complete cds.  
M34236.1 GI:335248  
M34236  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
FEATURES  
source  
CDS

TMGMPCPA 1413 bp ss-RNA linear VRL 03-AUG-1993  
Tobacco mild green mosaic virus movement and coat protein genes,  
complete cds.  
M34236.1 GI:335248  
M34236  
coat protein: movement protein.  
Tobacco mild green mosaic virus  
Tobacco mild green mosaic virus  
Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.  
1 (bases 1 to 1413)  
Nejdat, A., Cellier, F., Holt, C.A., Gafny, R., Eggenberger, A.L. and  
Beachy, R.N.  
Transfer of the movement protein gene between two tobamoviruses:  
influence on local lesion development  
Virology 180 (1), 318-326 (1991)  
91082424  
1984654  
Original source text: Tobacco mild green mosaic virus (strain PV  
228). cDNA to viral RNA.  
Draft entry and computer-readable sequence for [1] kindly submitted  
by A.Nejdat, 11-MAY-1990.  
Location/Qualifiers  
1. 1413  
/organism="Tobacco mild green mosaic virus"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:12241"  
51. 851  
/codon\_start=1  
/product="movement protein"  
/protein\_id="AAA47938.1"  
/db\_xref="GI:335249"  
/translation="MAVSLRDTYKISSEPTINLSKODELIRAFMTKYSVRISTDTKIMA  
VYNDSSLDVLLKGVLYKNGYCLGLVVGSMNLDKRGVSVICIDYKRRKRSNE  
ATUGAHAPACKKNSFSLIPNTYSIDAEKNPQVLYNIKGVAMEBGCPISLEFV  
SLCVHKNVKKGLRERILRTVDSPIELTEKVESEFVDEVPAAVLEFRRTKKKK  
RKKEKKRYVGVSNKKIINNSGKGLKVEIBDNVSDSEIASSTP"  
857. 1336  
/codon\_start=1  
/product="coat protein"  
/protein\_id="AAA47939.1"  
/db\_xref="GI:335250"  
/translation="MPYINSPOFVYLSAYADPVOLINLCTNALGNOFOTQOART  
VQOQZALDMKVPSPMTVRFPSADFYVRYNLTDLPLTALLNSPTTRKLTLEVDQPA  
PMTTEYNATORVDATVAIRASININLANELVGTGEMQAGETASGLVWTTTPAT"

Query Match 93.7%; Score 71.2; DB 14; Length 1413;  
Best Local Similarity 95.1%; Pred. No. 7.6e-14;  
Matches 73; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGACAGACGGCTGGCCAAATTGAACCTCACTGAAAAAGTTGTGAGAGCTTCGTGATGAA 60  
DB 551 GTGACAGACGGCTGGCCAAATTGAACCTCACTGAAAAAGTTGTGAGAGCTTCGTGATGAA 650  
QY 61 GTACCAATGGCTGTGA 76  
DB 651 GTACCAATGGCTGTGA 666

RESULT 5  
AB078435  
LOCUS  
DEFINITION  
Tobacco mild green mosaic virus complete genome, strain:Japanese.  
AB078435  
AB078435.1 GI:18253266  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
FEATURES  
source  
CDS

AB078435 6356 bp RNA linear VRL 22-JAN-2002  
Tobacco mild green mosaic virus complete genome, strain:Japanese.  
Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.  
1  
Okuno, T., Hamada, H., Takeuchi, S., Morishima, N., Yoshimoto, E. and  
Hikichi, Y.  
Nucleotide sequence of the Japanese isolate of Tobacco mild green  
mosaic virus  
Unpublished  
2 (bases 1 to 6356)  
Okuno, T., Hamada, H., Takeuchi, S., Morishima, N., Yoshimoto, E. and  
Hikichi, Y.  
Direct Submission  
Submitted (18-JAN-2002) Tetsuro Okuno, Kyoto University;  
Kitsashirakawa Oiwakecho, Sakyou 606-8502, Japan  
(E-mail:Okuno@kais.kyoto-u.ac.jp, Tel:81-75-753-6131,  
Fax:81-75-753-6131)  
Location/Qualifiers  
1. 6356  
/organism="Tobacco mild green mosaic virus"  
/mol\_type="genomic RNA"  
/strain="Japanese"  
/db\_xref="taxon:12241"  
/note="RNA transcripts synthesized in vitro from the cDNA  
clone of TMGMV-U direct systemic infection in Nicotiana  
benthiana-TMGWV-Japanese strain"  
72. 4901  
/note="position 3405..3407 is readthrough product"  
/codon\_start=1  
/trans\_except="(pos:3405..3407,aa:OTHER)  
/product="183kDa protein"  
/protein\_id="BA83986.1"  
/db\_xref="GI:18253267"  
/translation="MAHIOSTINALLSEVSGKNTLVNDLAPRMVDTADENAPAR  
RPKNPPTISEPOTLNTNAYREPOFTTPTNTQNAHSLAGLBALEETPLMCOVPG  
SPYVDGNFAHLFGKRDVHCMPULDIRHREGQDQSLMYIUSIRSNKVI  
EPORAAENRYAEAPNEVCCSTFOQDRHIPPENSGRYVALSLSLYDIPVHEGAAI  
SKOIHCVAFHFESEALLQTEVTLINEICATPKREBDDVSFFPADSTINMYHKN  
ILAVVKSYPFASRSRYVEKSPVATRNATPKCKTQDVTLYLKSXVGVGDSOFYE  
AMEDATYKALAMFETERALIDDTDSVNPFPKMKMVTPLPEGSITCKMKRSV  
IVNRDVTYVNIHRTYCAKALTYOVLSVSESIRSVIINGTARERMYDCAIIOP  
LSVPTPLQIKALLODDIYMKRCRCDKTSSELINDEVGKFGVGFITKELVSRKI  
LDVSENAIKIKIPDLVTWKDRFVAARYKSELPHIDIKOLEBAEQMYALSELSTL  
RADANFDIAFKDMCKALCVNPDVAARVIVAENNSGLTLPDCKPTEENVAALAKST  
ASEAVVLEPESDEVNKNESIAERGLPVCAASHGLTNANLEHOLESLNDFKACV  
DSVITKOMASVYVTSGLKVOQMKNYDVSALASISATVSNLCKSLKDVGVGDSRREK  
GVMDVTLKKWLLKBAKGSWGVLDYKGMFPALISYEDRWATESDMRRAVASSOT  
MYSDIATKONIKTRDSEPHPTAKMTIVDSVPGGKTKETLERDIDDEDLITVPG  
KQAAHITRRANSSGLIRATMDVARTVDSLELPKPRTRKRLTIDELMLHTGVNEL  
VLISGCDIAYVIGDQIIPINKVNPYPKPKHEKLOVDEEMKRTTLRGPGVNEFL  
OSKREBAVSTTSIVORSVSEMIKGVINSKPKLGIKIVTFEADKFELEKRYN  
VTVYHIOGETFEDSVLRITATPLTILSKSSSPHVALKTRITKSKFYTVVLDPLVQ  
IISDLSLSPLEMYNVEAGEXLOLQVAPVGHNLPAVATPSGPPLOLQVYDCL  
PNSITLKYDAVTWLRNDRNSINRQCVADFSSITPMKREKPCLEPVLRTAAEPVRY  
AGLENUVAMTRNRPAPDLTITIDESTASVAVDKEFDSYFLKSKYTKNIAGVTK

DSMMRLVSGKQERSTIGOLATNYFVDPIDQPKHMTIAKQPKQKIDLSIONEYPAALPTI  
VHSHQINIKIPEQPVSELTLLRLLEAVSDSPKLEFTRKAPDOPESFSDSHVMDVTV  
ELDIDSCQNRPHCAVEYEWKELGKINELPAFVWQKHQKXLLKYIKAIKTCIYAWY  
QKSDSDYTFEIGNTYIIAACLSGMLPMEKVIKGAFCGDSYLPYFKHIDPFSDISCAN  
LAKNREBAKIKRKGTFEGRIYIHHDKALITYDDKILISGKAHKKDYVHDELRER  
SLCDVACSIGNCAATFPOILAAIKRIVHKALIDDSFAPNCVNNFLCDKFLFRTLPLNGC  
72...3407

```

/codon_start=1
/protein="126kDa protein"
/protein_id="BAB83987.1"
/db_xref="GI:18253268"
/translation="MAHISTQTSNLLSEVSGKTLVNDLARRRMYDTAIDRFAARR
PKKNPPTCTISEBQTLITVNAYPEQILTFITQNAVHSLAGGLALSLRLSLKLVPG
STTDIGGFPAHLFKRGDVHCCMPNLDITIRHSDQKSLIMVSLISRLKRVPL
EFOEAPRFAVAPNRVCCSTQPDQRIHPENSGRRVYLSLHSLYVIYHFCFALAI
SKDIIICAAHPFSEBALLIDQETVLEIKIGATFPKRGDVSFPFADSTLVNFKRNV
ILHVYKSPFASRSIYFKEELVTRVWTFCKETKVDTYLLYKVSQVGDSDQFVE
AMEDEYFKKALAMENTERLEIPLDASVNFPEPKMDVILVPEGSI TSKMRKQV
LVNRFVTVTLNHRITQAKALYONVLSFESLRISVINGVATRAEMPDVIAILOD
LSMFEQTLTALLODDIYMKRERCLDTSESLIMDRYKRGVAPVPIRYSKVI
LDVSNNAKIKIPDLYVTWKDRPFAVYTKSEELPHLIDKDLSEAGQNTALSESLI
RDQANFEDLAKRDMCKALCVMPDGAARTVAENKSGTLITFKPKVRENAKALSTI
ASEAVRFLDPEVSDVMWNRKFSIARGKFLPVAEHSGLTINALEHQLSEINDIPKACV
DSVITQMAASVAYTGSILKVQGMKNYVSLASLSATVSNLCSTLKDVGVDVSDREKTV
GVMYDTLQKALIKPAAKHSWGVALDYKGRKFLTLISEGDRMDWRVAAVSQDT
MYVSIADQLQKRTKDSGEHEPTAKMVLVDVGQCGKTEILERVLDLGLITLVGK
KOAAAMIRRRANSGLIRATMDNRYTVSLFMBHPKPRHKLFLDEGLMHTGCGVNL
VLISCSQIAIYVGTQDQIPIINRYONPYPKHPKELVDREMRATLRLCGVNFEL
QSKKGAENSTTVSYORASVSEMGKAGKVLNPSPLQKGIYTFIQALQPELIEBNGYK
VNTYHIEIGETFEEDVSLVRLTATELITLISKSPVLAITRHTKSPKRYIVTLDPLQV
IISDLSLSSLSLMTMVEAGGR"
4891.5661

```

CDS

4891..5661  
/codon\_start=1  
/product="28kDa protein"  
/protein\_id="BAB33988.1"  
/db\_xref="gi:18253269"  
/translation="MAVASLRDTVKISERFDLSKQDDELPAPMTKVSVRISTDVKINA  
VKNSLSVDVLKGVKLVKNGYVCLAGLVSGENMLPDNCRGVSVCIYKRMKRSE  
ATLGAHAAPACKNPSFLLPNYSITSEDAEKHPQVLINIKGAMEGECPLSEV  
SICVYHKNRKGRLGRRLRLRTDGLPIELTEKVEEFVEDEPAAVXLERPRKTKRIV  
GNSVNNKKINNSGKKKGKVERIEDVSDEDSIASSSTP"  
5667..6146  
/codon\_start=1  
/product="17kDa protein"  
/protein\_id="BAB33989.1"  
/db\_xref="gi:18253270"  
/translation="MPYINSPSQFVYLSASAYDVPVLINLCTNALGNQFQTOQARTT  
VQQQFAAMKRVPSITVFPASDFYVRYNSTDELITALLNSFDTRKRLIETVNDQA  
PNTTEIVATKRVADATVAIRASINNLNDELVRCTGMNGAGETASGLTTPPA"

CDS

Query Match	91.6%	Score 69.6	DB 14	Length 6356
Best Local Similarity	94.7%	Fred No. 3.7e-13		
Matches 72	Conservative 0	Mismatches 4	Indels 0	Gaps 0
QY	1	GTGACAGCGGCTGTGGCCAACTGAATCAGTAAAGGTGTAGAGAGTTCGATGAA	60	
Db	5431	GTACACAGCGGCTGTGCCAATTGAATCACTGAAAGAGTGTAGAGAGTTCGATGAA	5490	
QY	61	GTACCAATGGCTGTGA	76	
Db	5491	GTACCAATGGCTGTGA	5506	

RESULT	6				
LOCUS	AX040184				
DEFINITION	Sequence 12 from Patent WO00633397.	411 bp	DNA	linear	EAT 18-NOV-2000
ACCESSION	AY040184				
VERSION	AX040184.1	GI:11230134			
KEYWORDS	.				
SOURCE	synthetic construct				
ORGANISM	synthetic construct				

mat\_peptide 1..909  
ORIGIN /product="cell-to-cell transport protein"

Query Match 60.3%; Score 45.8; DB 14; Length 912;  
Best Local Similarity 76.7%; Pred. No. 4.8e-05;  
Matches 56; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GTGACAGACGGCTGCCCAATTGAACCTACTGAAAAGTTGTGAGAGTTGTGATGAA 60  
DB 613 GTACACAGAGAGGGCCACCGCAACTTACTGAGCAGTTGTGATGATGATGAGAAA 672  
QY 61 GTACCAATGGCTG 73  
DB 673 GTTCTTATGGCTG 685

RESULT 8  
E03624  
LOCUS  
DEFINITION DNA encoding a part of complementary DNA to odontoglossum ringspot virus genome RNA.  
ACCESSION E03624  
VERSION E03624.1 GI:2171839  
KEYWORDS JP 1992144685-A/1.  
SOURCE Odontoglossum ringspot virus  
ORGANISM Odontoglossum ringspot virus  
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.  
1 (bases 1 to 5997)  
REFERENCE Isomura, Y., Matsumoto, Y., Chatani, M. and Ikegami, M.  
AUTHORS CDNA OF ORSV GENE  
TITLE Patent: JP 1992144685-A 1 19-MAY-1992;  
JOURNAL NIPPON OIL CO LTD

COMMENT OS Odontoglossum ringspot virus  
PN JP 1992144685-A/1  
PD 19-MAY-1992  
PF 28-NOV-1989 JP 1989306626  
PI ISOKURA YOSHIKATSU, MATSUMOTO YOSHITOMO, CHATANI MASAOKI, PI  
IKEGAMI MASATO  
PC C12N15/40, C07K13/00, C07K15/04, C12N1/21, C12N5/10, C12P21/02, PC  
(C12P21/02,  
PC C12R1:19);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: clone=PORB5-25;  
FH Key Location/Qualifiers  
FH misc\_feature 1..5997  
FT /note='a part of complementary DNA to FT  
FT Odontoglossum  
FT ringspot virus genomic RNA'  
FT /note='a part of complementary DNA to ORSV'.  
FEATURES  
source Location/Qualifiers  
1..5997  
/organism="Odontoglossum ringspot virus"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:12238"

ORIGIN

Query Match 60.3%; Score 45.8; DB 6; Length 5997;  
Best Local Similarity 76.7%; Pred. No. 7.1e-05;  
Matches 56; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GTGACAGACGGCTGCCCAATTGAACCTACTGAAAAGTTGTGAGAGTTGTGATGAA 60  
DB 4805 GTACACAGAGAGGGCCACCGCAACTTACTGAGCAGTTGTGATGATGATGAGAAA 4864  
QY 61 GTACCAATGGCTG 73  
DB 4865 GTTCTTATGGCTG 4877

RESULT 9  
E04305  
LOCUS  
DEFINITION cDNA encoding Odontoglossum ring spot virus (ORSV) genomic RNA.  
ACCESSION E04305  
VERSION E04305.1 GI:2172508  
KEYWORDS JP 1993030975-A/1.  
SOURCE Odontoglossum ringspot virus  
ORGANISM Odontoglossum ringspot virus  
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.  
1 (bases 1 to 6597)  
REFERENCE Isomura, Y., Matsumoto, Y., Chatani, M., Mizuta, Y. and Ikegami, M.  
AUTHORS CDNA OF ORSV GENE  
TITLE Patent: JP 1993030975-A 1 09-FEB-1993;  
JOURNAL NIPPON OIL CO LTD

COMMENT

OS Odontoglossum ring spot virus  
PN JP 1993030975-A/1  
PD 09-FEB-1993  
PF 26-JUL-1991 JP 1991276075  
PI ISOMURA YOSHIKATSU, MATSUMOTO YOSHITOMO, CHATANI MASAOKI, PI  
MIZUTA YOSHINORI, IKEGAMI MASATO  
PC C12N15/40, C07K15/04, C12N1/21, C12N15/11, C12N15/70, C12P21/02, PC  
C1201/68,  
PC C12N1/21, C12R1:19), (C12N15/70, C12N1:19), (C12P21/02, C12R1:19);  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: clone=PORCP-D83;  
FH Key Location/Qualifiers  
FH misc\_feature 1..6597  
FT /note='cDNA to Odontoglossum ring spot virus  
FT genomic RNA'  
FT CDS 5707..6183  
FT FT  
FT mat\_peptide /product='coat protein'  
FT /product='coat protein'.  
FEATURES  
source Location/Qualifiers  
1..6597  
/organism="Odontoglossum ringspot virus"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:12238"

ORIGIN

Query Match 60.3%; Score 45.8; DB 6; Length 6597;  
Best Local Similarity 76.7%; Pred. No. 7.3e-05;  
Matches 56; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GTGACAGACGGCTGCCCAATTGAACCTACTGAAAAGTTGTGAGAGTTGTGATGAA 60  
DB 5405 GTACACAGAGAGGGCCACCGCAACTTACTGAGCAGTTGTGATGATGATGAGAAA 5464  
QY 61 GTACCAATGGCTG 73  
DB 5465 GTTCTTATGGCTG 5477

RESULT 10  
ORU34586  
LOCUS  
DEFINITION Odontoglossum ringspot virus, complete genome.  
ACCESSION U34586  
VERSION U34586.1 GI:1407591  
KEYWORDS

ORGANISM Odontoglossum ringspot virus  
SOURCE Odontoglossum ringspot virus  
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.  
1 (bases 1 to 6609)  
REFERENCE Chung, C.G., Wong, S.M., Mahtani, P.H., Loh, C.S., Goh, C.J., Kao, M.C.,  
AUTHORS Chung, M.C. and Watanabe, Y.  
TITLE The complete sequence of a Singapore isolate of odontoglossum

JOURNAL  
MEDLINE  
PUBMED  
REFERENCES  
AUTHORS  
JOURNAL  
TITLE  
Submitted (22-AUG-1995) Sek-Man Wong, Botany, NUS, Kent Ridge,  
Singapore, 0511, Singapore

FEATURES  
Source

CDS

location/Qualifiers  
1.6609  
/organism="Odontoglossum ringspot virus"  
/mol\_type="genomic RNA"  
/strain="Singapore 1"  
/db\_xref="taxon:12238"  
join(63..3398,3402..4901)  
/codon\_start=1  
/product="183 kDa readthrough protein"  
/protein\_id="AAC55012.1"  
/translation="MAHFOOTMNNKVIIEAGMGRNSLINDLAQRVYDVAEELNHSR  
RPRVNSKVISOEOLIOATNAYPEFETPTNTQLAHVSAGGLRALLEYLMQIPG  
SITTDIGNFSALHYKGRDYHCMPENLIRVABRHINODTSTYLALERSKRLP  
VQGSARNTKNDPDAVCCDKRFQDCSYSGAGLPKRTYAGLSHYIIPADERAALR  
KDHICITAFHSENLILSTSAVLDEIGATFYKSDRLSFFQNSTLNYESHYTA  
IKYVCTFFPASRPFYHKEFMCTRVNMFCEKTKVDYFLFRGYTRGDSQPYTA  
MEAMEYKKTILAMNSERTIFRDAVNVNMFVKNMVIPLFDGVTGKMKRSVM  
VKDFEYVTLNIRTYODKALTYKNVLSVESISRSVIINGVARSMDVDSVLQAL  
SWPTLQTLAKRQDQVYKKQKEDDTYTNLFMKQISAPALPISIEBLISGFV  
KVAESIOIKTPDEYTPADKLVMYEQATEBLOHLDISPRABRYVALSELVTK  
EDBERDITPFAKCEKDIAPDYVAVYIPIMKNETLTPPKNTPALSDALSPLPKD  
LDMRFLKLTSCAPPSVKTLDGSLPKQSGYDERQESQSVSDPHLSVESVK  
NKSMSAVYTPGLKVOQMKNMIDYLSASISATVSNLCKVLKCGMRVAVASDSPIYSD  
VYKGMILKPKOKCHAMGVALNNGEKVYLLEMDAGPFCGMRVAVASDSPIYSD  
MGLTILSLCLDGEPLVLRMPKTYLDVGGGCKTLEITVNFEDLILVPGKACK  
MIRKANSGHVATREDNVRTYDSEFLMHLKPTNYNLFIDEGMLHTGCVNPLALSH  
CREAVRGGABEOLIPTRNANPRPKHRYTCLYHREVRSLRCADYTHRNKSKTD  
GELCTNDYIRSVDAEAVNGKVFNPKSPKLGKITTQSDALKEKGEYEVSTFG  
ELNVTHEIIGETFEVSVRLTPTELELSKSPHLVALTRHTSKFYKSVLDELV  
KWSDLISKVDFILDMYKVDAGILQVQVSIKGENLPFCERKSGVSIIMQYDYL  
PENSTILNEDAVTNMLRNNLNVNKDCITDFSKSVSPROQEFPPARTAAERPS  
AGLENLVAMIKRNPSPDLTGILIEDIATVNNKFMQVYIIDEISGNNTPMSDA  
FHRMAKOKKSTIGOLADPDYDLADIIOYKEMIKAKPQKGLISQDDEYALQTTVY  
HSKQINATFGLFSELTROLLERIDSKSLPTTRKTPBQIEAFSOLDSVPEBVL  
DISKTDQNEHFCAVEYIMERLGLNGLEVMQGRKSLKDYTAIKCLMWQR  
KSGDVTTPIGNVTYIAACLSMIPMDKVLKAFCCGDSITLYPKGLDPLDIOGNNR  
WNFEAKLYRKYGYFCGRYIIRHDKALVYDPLKLSLGGCHKLSLHLEFRSL  
CVSSSLNNCAVFGOLNDAIEVHKTAVNGSAPFCSIVYVLDXNLFRLFNNGSSTK  
G"

CDS

63..3431  
/codon\_start=1  
/product="126 kDa replicase"  
/protein\_id="AAC55013.1"  
/db\_xref="GI:1407593"  
/translation="MAHFOOTMNNKVIIEAGMGRNSLINDLAQRVYDVAEELNHSR  
RPRVNSKVISOEOLIOATNAYPEFETPTNTQLAHVSAGGLRALLEYLMQIPG  
SITTDIGNFSALHYKGRDYHCMPENLIRVABRHINODTSTYLALERSKRLP  
VQGSARNTKNDPDAVCCDKRFQDCSYSGAGLPKRTYAGLSHYIIPADERAALR  
KDHICITAFHSENLILSTSAVLDEIGATFYKSDRLSFFQNSTLNYESHYTA  
IKYVCTFFPASRPFYHKEFMCTRVNMFCEKTKVDYFLFRGYTRGDSQPYTA  
MEAMEYKKTILAMNSERTIFRDAVNVNMFVKNMVIPLFDGVTGKMKRSVM  
VKDFEYVTLNIRTYODKALTYKNVLSVESISRSVIINGVARSMDVDSVLQAL  
SWPTLQTLAKRQDQVYKKQKEDDTYTNLFMKQISAPALPISIEBLISGFV  
KVAESIOIKTPDEYTPADKLVMYEQATEBLOHLDISPRABRYVALSELVTK  
EDBERDITPFAKCEKDIAPDYVAVYIPIMKNETLTPPKNTPALSDALSPLPKD  
LDMRFLKLTSCAPPSVKTLDGSLPKQSGYDERQESQSVSDPHLSVESVK  
NKSMSAVYTPGLKVOQMKNMIDYLSASISATVSNLCKVLKCGMRVAVASDSPIYSD  
VYKGMILKPKOKCHAMGVALNNGEKVYLLEMDAGPFCGMRVAVASDSPIYSD  
MGLTILSLCLDGEPLVLRMPKTYLDVGGGCKTLEITVNFEDLILVPGKACK  
MIRKANSGHVATREDNVRTYDSEFLMHLKPTNYNLFIDEGMLHTGCVNPLALSH  
CREAVRGGABEOLIPTRNANPRPKHRYTCLYHREVRSLRCADYTHRNKSKTD  
GELCTNDYIRSVDAEAVNGKVFNPKSPKLGKITTQSDALKEKGEYEVSTFG  
ELNVTHEIIGETFEVSVRLTPTELELSKSPHLVALTRHTSKFYKSVLDELV  
KWSDLISKVDFILDMYKVDAGILQVQVSIKGENLPFCERKSGVSIIMQYDYL  
PENSTILNEDAVTNMLRNNLNVNKDCITDFSKSVSPROQEFPPARTAAERPS  
AGLENLVAMIKRNPSPDLTGILIEDIATVNNKFMQVYIIDEISGNNTPMSDA  
FHRMAKOKKSTIGOLADPDYDLADIIOYKEMIKAKPQKGLISQDDEYALQTTVY  
HSKQINATFGLFSELTROLLERIDSKSLPTTRKTPBQIEAFSOLDSVPEBVL  
DISKTDQNEHFCAVEYIMERLGLNGLEVMQGRKSLKDYTAIKCLMWQR  
KSGDVTTPIGNVTYIAACLSMIPMDKVLKAFCCGDSITLYPKGLDPLDIOGNNR  
WNFEAKLYRKYGYFCGRYIIRHDKALVYDPLKLSLGGCHKLSLHLEFRSL  
CVSSSLNNCAVFGOLNDAIEVHKTAVNGSAPFCSIVYVLDXNLFRLFNNGSSTK  
G"

CDS

KWSDLISKVDFILDMYKVDAGIL"  
4807..5718  
/codon\_start=1  
/product="movement protein"  
/protein\_id="AAC55014.1"  
/db\_xref="GI:1407594"  
/translation="MGLRFFVYLSIFPIKTEPSCSTYALVIRDSIKISEFNLSAS  
EKLPSALAVKSVIRISKVDLISYNDLSDLDLKGVLVNGVYCLAGLVATGBM  
NLDPNKGAVSICLNDKMRKABEATLGATYSACKRFEFKIIPYVSTADAKGI  
KQYMINIRGVBEMKGRCPISLBEPVSCVYIANTIKGLREKILNTEGTELETA  
DEFEKVPMAARLKSFRSVNKKRSDSKFVNKSKSLNKRKLNTEGSDVGLSVVD  
DIVVGVSDIRLDDCESFDQSDPY"  
5721..6197  
/codon\_start=1  
/product="coat protein"  
/protein\_id="AAC55015.1"  
/db\_xref="GI:1407595"  
/translation="MSYTTIDPSKLAIVSSAMDPNSLNLCTNSIGNOFQOQART  
VQOQFADVMPVPTLTSRFPAGGYRVVAYDILPLITFTLGTEDTRNRIEVENP  
QNETTETLDATEVDATAVIRSAINNLNLEIVRGVNGVNSFETMSGLTWTSS"

CDS

## ORIGIN

Query Match 60.3%; Score 45.8; DB 14; Length 6609;  
Best Local Similarity 76.7%; Pred. No. 7.3e-05;  
Matches 56; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
Cy 1 GTGACAGACGGCTGCGCAATTGACACCTGAAAGGTTGTGAGAGTTCGATGAA 60  
Db 5419 GTACAGAGAGAGGCGCCACCACTTCTGAGACAGTTTGTATGATGAGTTCGAGAGAA 5478  
Cy 61 GTACCAATGCTG 73  
Db 5479 GTTCTATGCTG 5491

## RESULT 11

LOCUS 583257 6611 bp RNA linear VRL 12-MAR-1997  
DEFINITION 126 kDa protein. .18 kDa protein [Odontoglossum ringspot virus,  
Cy-1, Genomic RNA, 4 genes, 66.1 nt].  
ACCESSION 583257  
VERSION 583257.1 GI:1881785  
KEYWORDS  
SOURCE Odontoglossum ringspot virus  
ORGANISM Odontoglossum ringspot virus  
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.  
REFERENCE 1 (bases 1 to 6611)  
AUTHORS Ikegami, M., Isumura, Y., Matsumoto, Y., Chatani, M. and Inouye, N.  
TITLE The complete nucleotide sequence of odontoglossum ringspot virus  
(Cy-1 strain) genomic RNA  
JOURNAL Microbiol. Immunol. 39 (12), 995-1001 (1995)  
MEDLINE 96381046  
PUBMED 8789059  
REMARK  
entry [NCBI bibs 179070] from the original journal article.  
Location/Qualifiers  
1..6611  
/organism="Odontoglossum ringspot virus"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:12238"  
66..4901  
/gene="183 kDa protein"  
join(66..3398,3402..4901)  
/gene="183 kDa protein"  
/codon\_start=1  
/protein\_id="AAB49498.1"  
/db\_xref="GI:1881786"  
/translation="MAHFOOTMNNKVIIEAGMGRNSLINDLAQRVYDVAEELNHSR  
RPRVNSKVISOEOLIOATNAYPEFETPTNTQLAHVSAGGLRALLEYLMQIPG  
SITTDIGNFSALHYKGRDYHCMPENLIRVABRHINODTSTYLALERSKRLP  
VQGSARNTKNDPDAVCCDKRFQDCSYSGAGLPKRTYAGLSHYIIPADERAALR  
KDHICITAFHSENLILSTSAVLDEIGATFYKSDRLSFFQNSTLNYESHYTA  
IKYVCTFFPASRPFYHKEFMCTRVNMFCEKTKVDYFLFRGYTRGDSQPYTA  
MEAMEYKKTILAMNSERTIFRDAVNVNMFVKNMVIPLFDGVTGKMKRSVM  
VKDFEYVTLNIRTYODKALTYKNVLSVESISRSVIINGVARSMDVDSVLQAL  
SWPTLQTLAKRQDQVYKKQKEDDTYTNLFMKQISAPALPISIEBLISGFV  
KVAESIOIKTPDEYTPADKLVMYEQATEBLOHLDISPRABRYVALSELVTK  
EDBERDITPFAKCEKDIAPDYVAVYIPIMKNETLTPPKNTPALSDALSPLPKD  
LDMRFLKLTSCAPPSVKTLDGSLPKQSGYDERQESQSVSDPHLSVESVK  
NKSMSAVYTPGLKVOQMKNMIDYLSASISATVSNLCKVLKCGMRVAVASDSPIYSD  
VYKGMILKPKOKCHAMGVALNNGEKVYLLEMDAGPFCGMRVAVASDSPIYSD  
MGLTILSLCLDGEPLVLRMPKTYLDVGGGCKTLEITVNFEDLILVPGKACK  
MIRKANSGHVATREDNVRTYDSEFLMHLKPTNYNLFIDEGMLHTGCVNPLALSH  
CREAVRGGABEOLIPTRNANPRPKHRYTCLYHREVRSLRCADYTHRNKSKTD  
GELCTNDYIRSVDAEAVNGKVFNPKSPKLGKITTQSDALKEKGEYEVSTFG  
ELNVTHEIIGETFEVSVRLTPTELELSKSPHLVALTRHTSKFYKSVLDELV  
KWSDLISKVDFILDMYKVDAGILQVQVSIKGENLPFCERKSGVSIIMQYDYL  
PENSTILNEDAVTNMLRNNLNVNKDCITDFSKSVSPROQEFPPARTAAERPS  
AGLENLVAMIKRNPSPDLTGILIEDIATVNNKFMQVYIIDEISGNNTPMSDA  
FHRMAKOKKSTIGOLADPDYDLADIIOYKEMIKAKPQKGLISQDDEYALQTTVY  
HSKQINATFGLFSELTROLLERIDSKSLPTTRKTPBQIEAFSOLDSVPEBVL  
DISKTDQNEHFCAVEYIMERLGLNGLEVMQGRKSLKDYTAIKCLMWQR  
KSGDVTTPIGNVTYIAACLSMIPMDKVLKAFCCGDSITLYPKGLDPLDIOGNNR  
WNFEAKLYRKYGYFCGRYIIRHDKALVYDPLKLSLGGCHKLSLHLEFRSL  
CVSSSLNNCAVFGOLNDAIEVHKTAVNGSAPFCSIVYVLDXNLFRLFNNGSSTK  
G"

## FEATURES

gene

CDS



SAGLENTLVAMIKRNENSPDLTGILDIEDTAEIVANKFMDAYIIDEISGNTYPTMTSD  
 AFHNNAKOEKSTIGQLADPEFDVLDPAIQYKIMIAQOKLIDPDEYALQITIV  
 YHSKOINAFEGPSELFTROLERIDSSSELTFTRTPEOIEFSPDSTVMEVLE  
 LDISKYDSONEHEKAVEYLIMERKLGNGLEWMOGRKISIKITAGITKICLMO  
 RASGDYTTIGTAVITAAACIASIMPMOKITAKFCGDDSIITYPKLIDPDIOSGNT  
 MNFENKLRKAYIGFCGYIITHDRGALVYDPLIKSLGKIKHISLDHLEEFIS  
 LCDVSSILNNCAVFGQLNDALAEVHTAVANGSPAFCSIVKYLSDKNLFTLTPYNGSST  
 KS"

CDS

64..3402

/gene="RDRP"

/codon\_start=1

/product="RNA-dependent RNA polymerase/replicase"

/protein\_id="AA587223.1"

/db\_xref="GI:46309860"

/translation="MAHFOQTNNKVIIEAGMNSLINDLAORRYVNAVEELNHRSR  
 RPKVNSKVISOEIOITQNAVPEEITFTYNTOLAHSNAGLRLELEYLMQIPEG  
 SYTVIGKNSALIKGRDYHACMNDLIRYARLINOODYSTVYLSKRLD  
 VFOQAFNKTRMDPDVCCCKRPFDCSYVDLPKTYAALHSITDIPDEGRALNR  
 KDVAHCYAAHSENLLETTSAPLDEISATYKSDRLSPFQNESTLNEYXNV  
 IKVCKTEPPASNRFLYHKEPMCTRVNTPCFKVDYFLRGVYTRGDEQPYTA  
 MDEAMEYKTLMLNSERTIPDRANVMPFKVMVIVPLFDSVTSGRKRSBYM  
 VKDFFVYTLNHTYODKALTYKVLSEVESIRSVIINGVARSMDVDSVLQAL  
 SMTFLQTLGAADQVVLKFKQEDDTVTNLFWQISDAVDLPSITKRLISGFFV  
 KYABOSIQKTPDEXITTPADKLMEYKATEELQHLDISPLRABRYVNALESSTIK  
 ECDERDITPFAKLCERKIDPPVYAKVYAIKMKELTLPKPTPEALSPLPKD  
 LDMRPDLKLSLTCAPPSPVKTLDSGLPKQSYGDERQFSQSVSDHLASVBSYK  
 MNMSAAVYTGPKVQOMKNVMDYASISATVSNCKVLQVYADPSASRSKGYD  
 VYKGMKLPKDKHGAVALNGBKIVILEMADGPIICGMWRVAVASDLSIYSD  
 MGKLOTLSCLKDGPVSDAVYTLVDGVRGKTEILETVNDEDLILVQKCKCK  
 MTKRANSGHVATKDNVRYTVDSPFLMKPTVNTLFTDEGLMHTGQVNTPLASH  
 CREANVRGDTBOIPIINRYANFPYKHPRTLYVDHREVRSLKCPADVTHPMNSKITD  
 GVULTSDVIRSVDAEAVRQGVFNPKSPKLGKIIITQSDABELKEGVEVSTEG  
 EVNATHEIIGETFEDEVSVRLTPTPELISKSPPHVALVTRHTYSFKYYSVVDPLV  
 KYCSPLSKVSDFLIMRYKVDAGIL"

gene

4808..5719

/gene="WP"

4808..5719

/gene="WP"

/codon\_start=1

/product="movement protein"

/protein\_id="AA587225.1"

/db\_xref="GI:46309862"

/translation="MGRILRFVYLISFPKITSSEPCSTMALVLRDSIKISEPINLSAS  
 EKLPSALTAVKSVRIKVDKTIYENDLSDIDLKGVLYVNGVYCLAGLVVIGEM  
 NLPDNCKGVSICLVDRMKRANEAITLSYHSACKORFTFKIIPNYSTYADALGI  
 MOVMTNIRGVEWKEGFCPLSEFVASICIYLNINIKLGLEKILNIVEGGPTLEAVV  
 DEFEKVPMAARLISFRSVNKKPSNMSKFNAGKSLRNKLNKLYENGSDVGTSVVD  
 DIVVNSGVSDIIRIDDCBSFDMQSDSY"

gene

5722..6198

/gene="CP"

5722..6198

/gene="CP"

/codon\_start=1

/product="coat protein"

/protein\_id="AA587226.1"

/db\_xref="GI:46309863"

/translation="MSYITTPDSKLAIVSSANADPNSLINLCTNSIGNGPOTQOARTT  
 VQOQFADWQPVLTLSRFPAGYFRVRYDPLDPLITFMGTFENRRIIEVNP  
 ONPTTETLIDATRVDDATVAIRSAINMLNELVGTGYNQVVSFEFMGLMTS"

ORIGIN

Query Match 60.3%; Score 45.8; DB 14; Length 6612;  
 Best Local Similarity 76.7%; Pred. No. 7.3e-05;  
 Matches 56; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 GTGACAGAGCGCTCGCCATTTGACTCAGTGAAGAGTTGTGAGAGTTGCGATGAA 60  
 DB 5420 GTTAAAGAGAGAGGCGCCACCGACTTATGAAAGAGTTGTGATTCGAGAGAA 5479  
 QY 61 GTACCAATGGCTG 73  
 DB 5480 GTTCTATGGCTG 5492

RESULT 13  
 LOCUS TOTMV2  
 DEFINITION TMV 3' end.  
 ACCESSION V01406  
 VERSION V01406.1 GI:62119  
 KEYWORDS coat protein.  
 SOURCE Tobacco mosaic virus  
 ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.  
 REFERENCE 1 (bases 1 to 1004)  
 AUTHORS Guillely,H., Jonard,G., Kukla,B. and Richards,K.E.  
 TITLE Sequence of 1000 nucleotides at the 3' end of tobacco mosaic virus RNA  
 JOURNAL Nucleic Acids Res. 6 (4), 1287-1308 (1979)  
 MEDLINE 79201054  
 PUBMED 109810  
 COMMENT KST TMV2.  
 FEATURES  
 source Location/Qualifiers  
 1..1004  
 /organism="Tobacco mosaic virus"  
 /mol\_type="genomic RNA"  
 /db\_xref="taxon:12242"  
 321..800  
 /note="unnamed protein product; reading frame (coat protein)"  
 /codon\_start=1  
 /protein\_id="CAA24685.1"  
 /db\_xref="GI:62120"  
 /db\_xref="GON:P03570"  
 /db\_xref="Swiss-Prot:P03570"  
 /translation="MSYITTPDSQVFLSSANADPIELNLCTNALGNOFOTQOARTV  
 VORSEVWKPSPQVTVRPDSDFKYRYNAVALDPLVTALLGAFDRNRIIEVNOAN  
 PTTAETLIDATRVDDATVAIRSAINMLNELVIRGTGSYNRSSPSSSGIWTSGPAT"

ORIGIN

Query Match 57.9%; Score 44; DB 14; Length 1004;  
 Best Local Similarity 73.7%; Pred. No. 0.00021;  
 Matches 56; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1 GTGACAGAGCGCTCGCCATTTGACTCAGTGAAGAGTTGTGAGAGTTGCGATGAA 60  
 DB 52 GTGACAGAGCGAGCGCCACCGACTTATGAAAGAGTTGTGATTCGAGAGAA 111  
 QY 61 GTACCAATGGCTG 76  
 DB 112 GTCCCTATGCAATCA 127

RESULT 14  
 LOCUS E00088  
 DEFINITION Genomic RNA containing gene of capsid protein.  
 ACCESSION E00088  
 VERSION E00088.1 GI:5708407  
 KEYWORDS UP 1983051894-A/2.  
 SOURCE Tobacco mosaic virus  
 ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.  
 REFERENCE 1 (bases 1 to 1019)  
 AUTHORS Roorensu,I.P. and Merit,K.H.  
 TITLE RNA PLANT VIRUS VECTORS OR FABRICATION AND UTILIZATION OF PART  
 JOURNAL NATL RES KAUINSURU OBU CANADA  
 COMMENT OS tobacco mosaic virus  
 PN UP 1983051894-A/2  
 PD 26-MAR-1983  
 PE 27-MAY-1982 JP 1982090482  
 PR 27-MAY-1981 US 81 267539  
 PT ROORENSU II PERUCHIYAA, MERIT KURISTEIN HARASA PC  
 C12N15/00,A01H1/00,C07H21/04,C12P19/34,C12P21/00//C12RI/91; CC  
 strandedness: Single;



CC topology: linear;  
CC hypothetical: No;  
CC anti-sense: No;  
FH Key location/Qualifiers  
FT 5'UTR 1..320  
FT misc\_feature 61..234 /note="region essential for formation of FT  
FT misc\_feature 235..320  
FT 321..749 /note="region for regulation" FT CDS  
FT 3'UTR /product="capsid protein"  
FT 750..1022.  
FEATURES  
source location/Qualifiers  
1..1019  
/organism="Tobacco mosaic virus"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:12242"  
ORIGIN  
Query Match 57.9%; Score 44; DB 6; Length 1019;  
Best Local Similarity 73.7%; Pred. No. 0.00021;  
Matches 56; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
QY 1 GTGACGAGCGGCTCGGCAATGAACTGACCTGAGAAAGGTTGTAAGAGTTGATGAA 60  
DB 52 GTGAGGAGCGAGGCGGCGCAATGAACTGACGAGAAAGGTTGATGATGATGAGAAAGAT 111  
QY 61 GTACCAATGGCTGTGA 76  
DB 112 GTGCCATGTCAATCA 127  
RESULT 15  
LOCUS D63809 6395 bp RNA linear VRL 13-FEB-1999  
DEFINITION Tobacco mosaic virus genomic RNA for 130K protein, 180K protein,  
30K protein and coat protein, complete sequence.  
ACCESSION D63809  
VERSION D63809.1 GI:1619995  
KEYWORDS coat protein; 30K protein; 180K protein; 130K protein.  
SOURCE Tobacco mosaic virus  
ORGANISM Tobacco mosaic virus  
REFERENCE  
1 (bases 1 to 6395)  
Viruses; ssRNA positive-strand viruses; no DNA stage; Tobamovirus.  
AUTHORS Chen, J., Matanabe, Y., Sako, N., Ohshima, K. and Okada, Y.  
TITLE Complete nucleotide sequence and synthesis of infectious in vitro  
transcripts from a full-length cDNA clone of a rakkyo strain of  
tobacco mosaic virus  
JOURNAL Arch. Virol. 141 (5), 885-900 (1996)  
MEDLINE 96265021  
PUBMED 8678834  
REFERENCES 2 (bases 1 to 6395)  
Sako, N.  
TITLE Direct Submission  
JOURNAL Submitted (07-AUG-1995) Nobumichi Sako, Faculty of Agriculture,  
Saga University, Laboratory of Plant Virology, 1 Honjo-machi, Saga,  
Saga 840, Japan (tel:0952-24-5191 (ex.2730), fax:0952-22-6274)  
FEATURES  
source location/Qualifiers  
1..6395  
/organism="Tobacco mosaic virus"  
/mol\_type="genomic RNA"  
/strain="Rakkyo"  
/db\_xref="taxon:12242"  
69..4919  
/note="position 3417..3419 is readthrough codon"  
/codon\_start=1  
/trans\_except="(pos:3417..3419,aa:OTHER)  
/product="180K protein"  
/protein\_id="BA09876.1"  
/db\_xref="GI:1619996"  
/translation="MAYTOTATTSALDTRGNNISVNDLAKRRLYDTAVDEFNARDR  
RPKVNFSKVISSEQTLATRAYPEFOITPVNTQNAVHSLAGLSLELEYLMQIPYG  
SLITDIGNFAELFKRAYVHCMPMLDVADRHRHGQKDSIELYSLRDLGGKTVP  
NPKAEADRAYPEFEDAVCHNTFCQCHOPMOQSGXYAIALSHIYDI PADPFGALL  
KXVHTCYAAPHSENLLEDSSVNTWFCRSDDKLTFSFASSTIYNCSFST  
ILKVCCTTPASRREYKMKFLVTRNTWFCRSDITFLIYGVNAKSDSDFPT  
AMEDAMHYKTLACNSERILLEDSSVNTWFCRSDITFLIYGVNAKSDSDFPT  
LVSKDFPHTVNTWFCRSDITFLIYGVNAKSDSDFPT  
LSMTFPLHTVNTWFCRSDITFLIYGVNAKSDSDFPT  
IRAGADLEIKVPDIYTFEDRLVAEKVSVDMPALDPRKMEETVYVNAKSDSDFPT  
RESGDPVDVFSOMCKLEFVDPMTAAVIAVWVNSGTLTPRPPEANAOALQD  
EKASGALVYTSRREYKMKFLVTRNTWFCRSDITFLIYGVNAKSDSDFPT  
TADSLRKQMSIYVTPITVQCKNTIDSLVSLVAVSLVILKOTAIIDETRO  
KFDVLDAKMSIYVTPITVQCKNTIDSLVSLVAVSLVILKOTAIIDETRO  
ESVYVSPMAKTLRLRLRDGEPHVNAKVLVGVGCGCTKGIILSRVNDDELITV  
PGKOAEMIRRRANSGLIYATKDNVTVVSPFMMNFGCKTFCORPLIDRGLMTG  
CNPVIVMSLDVAVVYVGTQOIPYINRVSGFPYAPAFSLKLENDVETRRPLACPD  
VTRHIVRRYSGFPVYVTSVYKVSQEWASGAAYNPISKLHGKILTFRTSGDKALLS  
KQYSEVHTVHGVGTSDVSLVRLIPTIISLAGDSPHVVALVSLRFTCSLKITVVM  
DPLVSLIRLEKLSYLLDMYKVDAGIQXQDQIDSVKSNLFPAAAPRTGDISMOPY  
YDCLPQNSNMANNFPAVWRLTDIISLVNDQCLDMKSVAAAPDOIKPLIPWRTAA  
EMPRQGLLENLVAIKRNPNABELGIDIENTASLVNKPFPSSYLKEKRNKRV  
SLFPRSINRWLEKROVITIGOLADPFDVPAVDYRHHIKAPKXKLDTSIGTEY  
ALOTIYVHSHKINAIFGPLEPSELTRDLSVDSRRPLFRTRKPAOIEDPFGDISHV  
PMVILHDSIKYKSGQNHFCAYBEIWRLEFEDPISGVWVKQHRITTLKDTYAGIK  
TCIYQKSGDVITFIGNITVLAQCLASMLRMEKIKGACGDSLLYPRGCEPDI  
QHSVNLMMNEEALFKKQYVFCGRYIHRDRCIYVYDPLKLSIKGAKIKDMEHL  
BEFRSLCDVAVSLNCAVYITQLDVWVEVHTAPDSFYKSLVYLSKVLFRSLF  
INSSC"  
69..3419  
/codon\_start=1  
/product="130K protein"  
/protein\_id="BA09877.1"  
/db\_xref="GI:1619997"  
/translation="MAYTOTATTSALDTRGNNISVNDLAKRRLYDTAVDEFNARDR  
RPKVNFSKVISSEQTLATRAYPEFOITPVNTQNAVHSLAGLSLELEYLMQIPYG  
SLITDIGNFAELFKRAYVHCMPMLDVADRHRHGQKDSIELYSLRDLGGKTVP  
NPKAEADRAYPEFEDAVCHNTFCQCHOPMOQSGXYAIALSHIYDI PADPFGALL  
KXVHTCYAAPHSENLLEDSSVNTWFCRSDDKLTFSFASSTIYNCSFST  
ILKVCCTTPASRREYKMKFLVTRNTWFCRSDITFLIYGVNAKSDSDFPT  
AMEDAMHYKTLACNSERILLEDSSVNTWFCRSDITFLIYGVNAKSDSDFPT  
LVSKDFPHTVNTWFCRSDITFLIYGVNAKSDSDFPT  
LSMTFPLHTVNTWFCRSDITFLIYGVNAKSDSDFPT  
IRAGADLEIKVPDIYTFEDRLVAEKVSVDMPALDPRKMEETVYVNAKSDSDFPT  
RESGDPVDVFSOMCKLEFVDPMTAAVIAVWVNSGTLTPRPPEANAOALQD  
EKASGALVYTSRREYKMKFLVTRNTWFCRSDITFLIYGVNAKSDSDFPT  
TADSLRKQMSIYVTPITVQCKNTIDSLVSLVAVSLVILKOTAIIDETRO  
KFDVLDAKMSIYVTPITVQCKNTIDSLVSLVAVSLVILKOTAIIDETRO  
ESVYVSPMAKTLRLRLRDGEPHVNAKVLVGVGCGCTKGIILSRVNDDELITV  
PGKOAEMIRRRANSGLIYATKDNVTVVSPFMMNFGCKTFCORPLIDRGLMTG  
CNPVIVMSLDVAVVYVGTQOIPYINRVSGFPYAPAFSLKLENDVETRRPLACPD  
VTRHIVRRYSGFPVYVTSVYKVSQEWASGAAYNPISKLHGKILTFRTSGDKALLS  
KQYSEVHTVHGVGTSDVSLVRLIPTIISLAGDSPHVVALVSLRFTCSLKITVVM  
DPLVSLIRLEKLSYLLDMYKVDAGIQXQDQIDSVKSNLFPAAAPRTGDISMOPY  
YDCLPQNSNMANNFPAVWRLTDIISLVNDQCLDMKSVAAAPDOIKPLIPWRTAA  
EMPRQGLLENLVAIKRNPNABELGIDIENTASLVNKPFPSSYLKEKRNKRV  
SLFPRSINRWLEKROVITIGOLADPFDVPAVDYRHHIKAPKXKLDTSIGTEY  
ALOTIYVHSHKINAIFGPLEPSELTRDLSVDSRRPLFRTRKPAOIEDPFGDISHV  
PMVILHDSIKYKSGQNHFCAYBEIWRLEFEDPISGVWVKQHRITTLKDTYAGIK  
TCIYQKSGDVITFIGNITVLAQCLASMLRMEKIKGACGDSLLYPRGCEPDI  
QHSVNLMMNEEALFKKQYVFCGRYIHRDRCIYVYDPLKLSIKGAKIKDMEHL  
BEFRSLCDVAVSLNCAVYITQLDVWVEVHTAPDSFYKSLVYLSKVLFRSLF  
INSSC"  
CDS  
69..4919  
/codon\_start=1  
/product="coat protein"  
/protein\_id="BA09879.1"  
/db\_xref="GI:1619999"  
/translation="MSYNTNTPSQVYFSSAAADPEILINLTNMGKQNPOTQOARTV  
VQKQSEVWKSPPQVTPRPDSDFKYNANAVLDPVTLALPAPTRNITIEVNGAN  
PSTAEITLADRVDATVAIRSAINNLIVELTRGTSYVRSSESSGVLWTSPPAT"

ORIGIN

CDS

CDS

CDS



